

102(b)
11(mouse)

```
US-09-327-7500-12 x BG088461
Align seg 1/1 to: BG088461 from: 1 to: 599

1 MetAlaAsnValHisGluGluAsnGluGluMetGluGluProLeuGlnAs 17
175 ATGGCCCAATGTCACCAAGAAACAGAGATGGAGAGCCCTCGAGAA 224
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGluProAlaG 34
225 TGGACAGAAAGACCGCCCTGGGGAGAGGTGGAGCCACACCGCTGCTG 274
34 LysAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArgArg 50
275 CAACACACACACACACACACACACACACACACACACACACACACACAA 324
51 GlyGlnAlaArgGluLeuAlaProAsnPheArgTrpAlaTleProAsnAr 67
325 GGCCAGGCTCGCCGACTTGGCCCTTAACCTTCCGATGGGGCATTCCCA 374
67 gClnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
375 GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
84 etGluGluMetArgGluLeuArgGlyGlyGlyGlyGlyGlyGlyGlyG 100
425 TGGAGAGATGAGAGAGATCGGAGAGAGCTTAGGAGCTACAGCTGAGA 474
101 AsnCysLeuArgGluLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
475 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
117 sAspGluPheCysLeuMetPro 124
525 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 546
seq_name: gb_est1:AA272375

seq documentation block:
LOCUS AA272375 612 bp mRNA EST 26-MAR-1997
DEFINITION vB62g07.r1 Barstead mouse pooled organs MRLB4 Mus musculus cDNA
clone IMAGE:761628 5', similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AA272375
VERSION AA272375.1 GI:1910706
KEYWORDS EST.
SOURCE mouse house.
ORGANISM Mus musculus
REFERENCE
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gaisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMNI Mouse EST Project
Unpublished (1996)
CONTACT: Marka M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:462548
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 507.
LOCATION/Qualifiers
1..612
/organism="Mus musculus"
```

```
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="Barstead mouse pooled organs MRLB4"
/seq_auth="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"
/notes="Organ: pooled; Vector: pTT73D-Pac (Pharmacia) with
a modified polylinker; Site:1: EcoRI; Site:2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTCAGAAATCTCAAGTGGAGCGGCCCTTTTTCCTTTTTCCTTTTTCCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
[GTGGATCGGTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified pTT73 vector.
Library constructed by Bob Barstead."
BASE COUNT 168 a 167 c 175 g 102 t
ORIGIN

alignment_scores:
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:
US-09-327-7500-12 x AA272375
Align seg 1/1 to: AA272375 from: 1 to: 612

1 MetAlaAsnValHisGluGluAsnGluGluMetGluGluProLeuGlnAs 17
175 ATGGCCCAATGTCACCAAGAAACAGAGATGGAGAGCCCTCGAGAA 225
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGluProAlaG 34
225 TGGACAGAAAGACCGCCCTGGGGAGAGGTGGAGCCACACCGCTGCTG 275
34 LysAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArgArg 50
275 CAACACACACACACACACACACACACACACACACACACACACACACAA 325
51 GlyGlnAlaArgGluLeuAlaProAsnPheArgTrpAlaTleProAsnAr 67
325 GGCCAGGCTCGCCGACTTGGCCCTTAACCTTCCGATGGGGCATTCCCA 375
67 gClnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
375 GCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
84 etGluGluMetArgGluLeuArgGlyGlyGlyGlyGlyGlyGlyGlyG 100
425 TGGAGAGATGAGAGAGATCGGAGAGAGCTTAGGAGCTACAGCTGAGA 475
101 AsnCysLeuArgGluLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
475 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525
117 sAspGluPheCysLeuMetPro 124
525 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 547
seq_name: gb_est2:W81757

seq documentation block:
LOCUS W81757 616 bp mRNA EST 12-SEP-1996
DEFINITION W81757 mouse embryo NMWELL 5 14.5 Mus musculus cDNA
clone IMAGE:403307 5', similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION W81757
VERSION W81757.1 GI:1392776
KEYWORDS EST.
SOURCE house mouse.
```


Mar 12 09:01:15 2002
Cms 1, 2, 3, 4, 11(mouse)¹³, 20, 21, 22, 23

13-APR-1999

237 a 212 c 228 g 177 t

1

Align seg 1/1 to: AF097440 from: 1 to: 854

	seq_documentation_block:	
LOCUS	Af187065	519 bp mRNA ROD 11-JUN-2000
DEFINITION	Rattus norvegicus p75NTR-associated cell death executor (Nade)	
	mRNA, complete cds.	
ACCESSION	Af187065	
VERSION		
KEYWORDS		
SOURCE	Af187065.1 GI:8452895	
ORGANISM	Norway rat.	
	Rattus norvegicus	
	Eukaryota; Metazoa;	
	Mammalia; Eutheria; Rodentia;	
	Carnivora; Muridae; Murinae,	
	Rattus	
REFERENCE	1 (bases 1 to 519)	
AUTHORS	Mukai,J., Hachiyama,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D., Suvanto,P., Hanaka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A. NADE, a p75NTR-associated cell death executor, is involved in signal transduction mediated by the common neurotrophin receptor p75NTR	
JOURNAL	J. Biol. Chem. 275 (23), 17566-17570 (2000)	
MEDLINE	20298829	
REFERENCE	2 (bases 1 to 519)	
AUTHORS	Hanaka,T., Hachiyama,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P., Mukai,J., Li,Y., Irie,S. and Sato,T.	
TITLE	Characterization of NADE, a novel member of the TNF superfamily, which mediates neuronal apoptosis via interaction with p75NTR	
JOURNAL	Neuroscience Letters 277-280 (1999) Ocularyngology/Pathology, Columbia University, 630 West 168th St., Pk5 11-451, New York, NY 10032, USA	
FEATURES	Location/Qualifiers	
SOURCE	1..519	

FEATURES	
source	Location/Qualifiers
1. 519	

Tue. 1

```

/strain="FVB/N"
/db_xref=taxon:10090"
/clone_image=761628"
/clone_lib="Barstead mouse pooled organs MRLB84"
/clone_method="cloned"
/tissue_tag="pooled organs"
/phase_tag="day"
/lab_host="DH10B"
/notes=organ: pooled; Vector: pMT3D-pac (Pharmacia) with
a modified polylinker; Site_1: EcoRI; site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACAGTCATGAAGTGGGAGCGGCCCTTTTTTTTTTTTTTTTTTTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(CGTGACTCATGGTACC), digested with Not I and cloned into
the BamHI and EcoRI sites of the modified pMT3 vector.
Library constructed by Dr. Barstead."
BASE COUNT      168 a   167 c   175 g   102 t
ORIGIN

alignment_scores:
    Quality:     688.00      Length:     124
    Ratio:       5.548      Gaps:         0
    Percent Similarity: 100.000    Percent Identity: 98.387

Alignment_block:
US-09-327-750D-12 x AA272375

Align seq 1/1 to: AA272375 from: 1 to: 612

1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
|||||
176 ATGCCCATGTGCACCAAGAGAAMAAAGAAAGATGGAGCAGCCCTGGAGAA 225
|||
17 nGlyGluGluAspArg-ProValGlyGlyGlyGluGlnProLeuAg 34
|||||
226 TGGACAGAAAGAACGCCCTGTGTGGAGAGAGGTGGAGGCCACCACTGCTG 275
|||
34 TyAanAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnGarg 50
|||||
276 CAACACACACACACACACACACACACACACACACACACACACACACACGAGA 325
|||
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTfPalaIleProAsnAr 67
|||||
326 GGCCAGGCTCGCCNCTGTCCCCTTACTTCGATCGGATGCCCATTTCCCAACAG 375
|||
67 gLinMetAsnAPgLyLeuGlyGlyAspGlyAspAspMetGluMetPheM 84
|||||
376 CGAGATGTAATGACGGGTTGGGTGGAGATGAGATGATGGAANAATTCCA 425
|||
84 etcGluMetArgGluIleArgArgTyLeuArgGluLeuGlnLeuArg 100
|||||
426 TGGAGAGATGAGAGATGCTGGGAAAGAGCTTAGGGAGCTACAGCTGAGA 475
|||
101 AsnCysLeuArgTlleLeuMetGlyGluLeuSerAsnHisHisAspHisHl 117
|||||
476 AATTGTCTACGATCCCTTATGGGGAGCTGTCTTACCACCCAGCATGCCA 525
|||
117 saspGluPheCysLeuMetPro 124
|||||
526 TGATGAATTCCTGCCATTATGCCCT 547

eg_name: gb_est2:N81757

eg_documentation_block:
LOCUS      N81757                616 bp      mRNA
DEFINITION me95dd06.r1 Soarax mouse clone gp N81757.5 EST
            clone IMAGE403307.3; similar to CD-148188.5 Mus musculus cDNA
            13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
CESSION    N81757
VERSION    N81757.1 GI:1192776
KEYWORDS   EST.
```

51 GlycInaIaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
52
53
54
55
56
57 GCCAGGCTGCGCGACTTGCCTTAAGTTCGATGGCCATATCCCAACG 206
58
59
60
61
62
63
64
65
66
67 gclnMcAsnAspGlyLeuGlyAspGlyAspAspMetGluMetPheX 84
68
69
70
71
72
73
74
75
76
77 CAGTGAATGAGCGGGTGTGGTGAGTGGAGATGATGGAAATGTCTCA 256
78
79
80
81
82
83
84 cTgGluMetArgGluIleArgArgCysLeuArgGluLeuGluLeuArg 100
85
86
87
88
89
90
91
92
93
94
95
96
97 TGGAGGAGATGATGACAGATCCGGGAGAAGCTTAGGGAGCTACACGTGAGA 306
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842

seq name: qb est1:AW476468

seq_name: qb_est1:AW476468

LOC603	nmr7602.y1	NCI_CGAP_Lu33	Mus musculus	cdna clone	IMAGE:2937027	5'
DEFINITION	uq76602.y1	NCI_CGAP_Lu33	Mus musculus	cdna clone	IMAGE:2937027	5'
	similar to qb:M38188	OVARIAN GRANULOSA CELL	13.0 KD PROTEIN HGR74			

ACCESSION
NUMBER
AM476458
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
house mouse,
Mus musculus

GI:7046574

(HUMAN); mRNA sequence.

COMPLETION DATE: 1990-09-10

REFERENCE 1 (bases 1 to 628)

1 (pages 1 to 240)
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS
 TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Published (1997)
 Unpublished (1997)
 JOURNAL

CONTACT: Robert Strausberg, Ph.D.
Unpublished (1997)
Email: csapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CONV. Tissue Procurement: M. Bostock

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center

For distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium, <http://imaging.nci.nih.gov/bbrp/image/image.htm>

NO. 47
 Supplier: -40RP from Gibco
 Hi quality sequence stop: 456.
 Location/Qualifiers
 1..528
 source

```

source
1. 020
/organism="Mus musculus"
/strain="C2ECH II"
/db xref="taxon:10090"

```

[illegible]

```

/organ="lung; Vector: pr773D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
/lab_host="DH10B (phage-resistant)"
/tissue_type="pooled lung tumors"
/organ="lung; Vector: pr773D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st

```

immobilized poly(amine), site₂, NOL1, site₂, EONA, 180
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAATGGGAGCGGGCCCTCTGTTTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "1 and 20 of sites of the modified pV13 vector. Library

constructed by Bento Soares and M. Fatima Bonafide.

(9) 201

```

BASE COUNT      162 a   157 g   172 g   93 t     1 others
ORIGIN
Alignment_scores:
    Quality: 688.00      Length: 124
    Ratio: 5.548        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 98.387

alignment_block:
US-09-327-7500-12 x BE334877 ..
Align seg 1/1 to: BE334877 from: 1 to: 595

1 MetAlaasnValHisGlnGluAsnGluMetGluGlnProLeuGlnAsp 17
166 ATGGCCCATGTGCCACAGCAAAACAAGAAGATGAGCAGCGCCCTGCAGAA 215
17 nclYgLuGluAspArgProValGlyGlyGlyGlyGluGlyHISLInPtoALag 34
216 TGGACAGACAGACCCTTCCTGTGGAGAGAGTGTAGGGCCACCAACCTGCTG 265
34 LysAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
266 CAACAACAACAACAACAACAACAACCAACCAACCAACCAACCAACCAAGA 315
51 GlyGlnAlaAarGargLeuAlaProAsnPheArgTrpAlaIleProAsnAr 57
316 GCCCAGGCTGCCGACTTGCCCTTAACCTTCGCATGGGGCCATTCCCAACAG 365
67 GClMetAsnAspGlyLeuGlyGlyAspGlyAspPheMetGluMetPheM 84
366 GCAGTGAAGACGGTTGGGTGGAGATGGAGATGGATGGAGATGGATTCGA 415
84 stGlUgluMetArgGluIleAargLyLeuArgGlyLeuGlnLeuArg 100
416 TGGAGGAGATGACAGAGATCCGGAGAAAGCTTAGGAGACTACAGCTGAGA 465
101 AancysLeuAarGileLeuMetGlyGlyLeuSerAsnHisHisAspHisH 117
466 ANTGTCTACCATCTTATGTGGGGAGACTGTCTAACCAACCAACCAACCA 415
117 AspGluPhuCysLeuMetPro 124
516 TCATGAATTCCTGCCTTATGCTT 537

seq_name: gb_est2:H64711

seq_documentation_block: 590 bp
NC_013111.1 | Homo sapiens chr11:54,538,181-54,538,188 |
DESCRIPTION
clone IMAGE:313872 / 54,538,181-54,538,188 | CHAFMAN
10 kb CD PROTEIN HGR74 (HUMAN) : mRNA sequence
EST

```

KEYWORDS
SOURCE
ORGANISM

est.
mouse moule:
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Sciurognathi; Muridae; Murinae; Mus
(bases 1 to 590)

REFERENCE
AUTHORS

Galecki,
Galecki,
Schellenberg,
Thieling,
Waterston,
Wyllie,
Lennon,
Soares,
Wilson,
Dieterich,
Martin,
Moore,
Underwood,
Moore,
Wilson,
R.

TITLE
JOURNAL
COMMENT

The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tue

/cgn2_6/ptodata/2/1ne/5B_COMB.seq:US-08-487-595-8 -	76.50	120.21	47.95	5077
/cgn2_5/ptodata/2/1ne/6A_COMB.seq:US-09-024-020B-8 -	76.50	117.25	70.14	432
/cgn2_5/ptodata/2/1ne/6A_COMB.seq:US-09-356-952-12 -	76.50	98.66	763.02	686
/cgn2_5/ptodata/2/1ne/5A_COMB.seq:US-08-415-611-47 -	76.00	134.65	7.53	108
/cgn2_5/ptodata/2/1ne/5A_COMB.seq:US-09-025-681-2 -	76.00	132.90	9.42	1294

```

1  seq_name: vcgn2_6/prodata7/1/ins/PCTUS_COMB.seq:PCT-US92-06840-1
2
3  seq_documentation block:
4  -----
5  Sequence 1 Application PC/TUS9206840
6  -----
7  GENERAL INFORMATION:
8  -----
9  APPLICANT: Shi, Yang
10 -----
11 APPLICANT: Seto, Edward
12 -----
13 APPLICANT: Shenk, Thomas
14 -----
15 TITLE OF INVENTION: Y11 TRANSCRIPTION FACTOR AND METHODS OF
16 -----
17 TITLE OF INVENTION: ISOLATING SAME
18 -----
19 NUMBER OF SEQUENCES: 10
20 -----
21 CORRESPONDING SEQUENCE:
22 -----
23 ADDRESSEE: Ostralenk, Faber, Gerb & Soffen
24 -----
25 STREET: 1180 Avenue of the Americas - 7th Floor
26 -----

```

CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-8403
 COMPANY: IBM
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent-In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/06840

```

1  APPLICATION NUMBER: PCT/US92/06840
2  FILING DATE: 19920814
3  CLASSIFICATION:
4  CLASSIFICATION: A61K 31/00
5  CLASSIFICATION: A61K 31/00
6  PRIOR APPLICATION DATA:
7  APPLICATION NUMBER: US 07/746,485
8  FILING DATE: 16-AUG-1991
9  ATTORNEY/AGENT INFORMATION:
10 NAME: Dennis, Manette
11 REGISTRATION NUMBER: 30,623
12 REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
13 TELECOMMUNICATION INFORMATION:
14 TELEPHONE: (01212) 3840700
15 TELEFAX: (01212) 3840688
16 TELEX: 2365925
17 INFORMATION FOR SEQ ID NO: 1:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 2353 base pairs
20 TYPE: NUCLEIC ACID
21 STRANDEDNESS: double
22 TOPOLOGY: linear
23 MOLECULE TYPE: DNA (genomic)
24 RECOMBINATION: NO
25 ANTI-SENSE: NO
26 ORIGINAL SOURCE:
27 ORGANISM: Homo sapiens
28 TISSUE TYPE: Hela cells derived from cervical
29 TISSUE TYPE: carcinoma
30 CELL TYPE: tumor cells
31 CELL LINE: Hela
32 IMMEDIATE SOURCE:
33 LIBRARY: D96/AH-2
34 CLONE:
35 FEATURE: p14-1 or p1Y1
36 NAME/KEY: CDS
37 LOCATION: 241..1485
38 PCT-US92-06840-1
39
40 alignment_scores:
41 Quality: 84.50
42 E-Value: 1.47E-05
43 Percent Siml: 98.00
44 Percent Ident: 98.00
45 Length: 866
46 Gaps: 28
47 Percent Var: 0.00

```

```
alignment_scores:
  Quality: 84.50
  Ratio: 1.798
  Percent Similarity: 54.65
  Percent Identity: 38.070
  Gaps: 3
  Length: 86
```

```

528 CGGCGACCGCGT...GAACATCGACCG 553
63 laile...ProteinArg...GlnMetAspGlyLeu...GlyGly 75
554 CAGTCACAGCGCTCGCTGACCGTCACGACGACGCTCGAAGCGGA 603
76 AspGlyAspMetGluMetPheMetGluGluMetArg... 88
604 GATCAACCATGATCAAGAGTCGCGGCGCCACGCGCGCTCGGA 653
89 .....GluileArgArgLysLeuA 95
654 TCGAGCGACCGCGGAGTTCACAAAGCTGATCTCGGCGCGCGACG 703
95 rgluLeuGlnLeuArgAncCysLeu...Argile..... 105
704 ATCTACTCCAGCGAGCGCTGCTCGCTCGCTCAAGCTCGCGACGG 753
105 ..... 105
754 CAGCACCTACTTCTGACCGCGCGCACTGACGCGGCGCGACCA 803
106 LeuMetGlyGluLeuSerPheHisAsp...HisHisAsp 118
804 CCGTGGCGGCACTCGCGCGCGCGCGCGCGCTCGCGACGAC 846
seq_name: /c9n2.5/ptodate/2/1na/6A_COMB.seq:US-09-171-969-1
seq_documentation_block:
: Sequence 1, Application US/09171969
: Patent No. 6284513
: GENERAL INFORMATION:
: APPLICANT: Thomas, Lawrence J.
: TITLE OF INVENTION: PLASMID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Banner & Witcoff, Ltd.
: STREET: 75 State Street, Suite 2300
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: FILING DATE: 01 May 1997 (01.05.97)
: APPLICATION NUMBER: US/09/171.969
: PRIORITY DATE: 01 May 1997 (01.05.97)
: CLASSIFICATION: 514
: PRIORITY NUMBER: 08/802.967
: FILING DATE: 01 May 1996 (01.05.96)
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/802.967
: FILING DATE: 21 February 1997 (21.02.97)
: ATTORNEY/AGENT INFORMATION:
: NAME: Leon R. Yanovich
: REGISTRATION NUMBER: 30,237
: REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1488 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL:
: ANTI-SENSE:
: FEATURE: Structural coding sequence for
: FEATURE: mature rabbit CETP

```

```

NAME/KEY:
LOCATION:
PUBLICATION INFORMATION:
AUTHORS: Nagashima, Mariko, et al.
TITLE: Cloning and mRNA tissue distribution of rabbit cholesteryl ester tr.
JOURNAL: J. Lipid Res.
VOLUME: 29
ISSUE: 1643 - 1649
PAGES: 1988
DATE: 1988
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1488
US-09-171-969-1

alignment_scores:
Quality: 81.00 Length: 101
Ratio: 1.373 Gaps: 3
Percent Similarity: 58.416 Percent Identity: 28.713

alignment_block:
US-09-327-750d-12 x US-09-171-969-1
Align seg 1/1 to: US-09-171-969-1 from: 1 to: 1488
5 H1SGInGluAsnGluMetGluGlnProLeuGlnAsnGlyGluGluAs 21
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
30 CATCGTGTGTGTCATCATCAAGCGCGCGCTCTT.....GGTGTGAACC 73
21 pArgProValGlyGlyGluGlyHisGlnProAlaGlyAsnAsnAsnA 38
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
74 AAGAGAGCGCGAAGTGTGTCACAGCGCGCTTCACGCGCGCGCGCTATCCG 123
38 snAsnAsnHisAsnHisAsnHisAsnHisAsnHisArgArgGlyGlnAlaArg 54
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
124 GAGCTCAGCGCGGAGAGCGCGCTGATGCTCTCGCGCGCGCTCAAGTACGG 173
55 ArgLeuAlaProAsnPheArgTTPAlaLeProAsnArgGlnMetAsnAs 71
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
174 GCTGCACACCTCCAGATCATCGACCGCGCTGC...CATCGCGACGCGCAGG 220
71 pGlyLeuGlyGlyAspGlyAspMetGluMetPheMetGluGluMetA 88
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
221 TCGAGCTGTGTGACGCGCAAGACCATCGACGCTCGCGCATCGCAAGACGTGTC 270
88 rgluileArgArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArg 104
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
271 GTGCTCTCAAGCGGAC...CCTGACTACACTACACGATGCTCTGGG 317
105 ile 105
318 GTT 320
seq_name: /c9n2.5/ptodate/2/1na/6A_COMB.seq:US-09-036-987A-1
seq_documentation_block:
: Sequence 1, Application US/09036987A
: Patent No. 6284513
: GENERAL INFORMATION:
: APPLICANT: Baliz, Richard H.
: APPLICANT: Broughton, Mary C.
: APPLICANT: Crawford, Kathryn P.
: APPLICANT: Madduri, Krishnamurthy
: APPLICANT: Merlo, Donald J.
: APPLICANT: Treadway, Patti J.
: APPLICANT: Turner, Jan R.
: APPLICANT: Waldron, Clive
: TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dow Agrosciences LLC Patent Department
: STREET: 9330 Zionsville Road
: CITY: Indianapolis

```

95. rgluLeuInLeu.....ArgAnCysLeu.....Atg 104
280623 ACAGCGCGACGTTTCGGCCATCATGTCGGTGGCGGACCAACGACGCG 280672
105 lleLeuMetGlyLeuSerAsnHisHis.....AspHis 116
280673 TTCACACAGGGGAAGTTCCTGATGCATCCACGCCGACGATCAC 280714
seq_name: /cgn2_6/prodata/2/lna/6B_COMB.seq:US-09-103-840A-1
seq_documentation_block:
Sequence 1. Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: FRASER, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: H3TRV
US-09-103-840A-1

alignment_scores:
Quality: 80.00 Length: 164
Ratio: 1.013 Gaps: 8
Percent Similarity: 48.171 Percent Identity: 23.171
US-09-327-7500-12 x US-09-103-840A-1
Align seg 1/1 to: US-09-103-840A-1 from: 1 to: 4411529
5 HisGInGluAsnGluMetGluInProLeuInAsn...GlyGluG 120
280110 CATCCGGCGGAGAAATTTGTGCAGCAAAATGCTGTAATATGTCGGGTGA 280159
20 uAspArgProValGlyGlyGlyGly.....HisGlnProA 33
280160 GATATACGCTTTCGGCGGATGCAGATGTCGGTGTGTCGACCCG 280209
33 lacGlyAsnAsnAsnAsnAsnHis.....AsnHisAsnHisAsn 46
280210 CCGGTGACGGCGGTGCTGACCGCGGCTCCAGATGTCGGCGCATAC 280259
47 HisHisArgArgGlyGlnAlaArg..... 54
280260 GATGACGGTGTGCTCCCGCGAGCTCGACAGCGTGCAGACGGTGTGATCTCT 280309
55ArgLeuAlaProAsnPheA 61
280310 TACCGCGGATAGCGCCACCGATTGGCGCGCGCTCGCTTCGGTCAGC 280359
61 rgtPralleProAsnArgGlnMetAsnAspGlyLeuGly..... 75
280360 GTGGCCCGCGGACCGCGGGATACAGCAGGATGCTTCGACGGCTCCGA 280409
76AspGlyAs 78
280410 GCTAACAGCACGCTCTGGAAGCAGCGCTCCGGAGACCGCTCGGGCA 280459
78 pAspMetGluMetPheMetGluLeuMetArgGluLeuArgGlyLeuA 95
280573 TGACGTGGCGAGCTACAGCGCGCATATTCGCGGACGTTTCGACGGCTGTTG 280622

alignment_block:
US-09-327-7500-12 x US-09-103-840A-2
Align seg 1/1 to: US-09-103-840A-2 from: 1 to: 4403765
5 HisGInGluAsnGluMetGluInProLeuInAsn...GlyGluG 120
280223 CATCCGGCGGAGAAATTTGTGCAGCAAAATGCTGTAATATGTCGGGTGA 280272
20 uAspArgProValGlyGlyGlyGly.....HisGlnProA 33
280273 GATATACGCTTTCGGCGGATGCAGATGTCGGTGTGTCGACCCG 280322
33 lacGlyAsnAsnAsnAsnAsnHis.....AsnHisAsnHisAsn 46
280323 CCGGTGACGGCGGTGCTGACCGCGGCTCCAGATGTCGGCGCATAC 280372
47 HisHisArgArgGlyGlnAlaArg..... 54
280373 GATGACGGTGTGCTCCCGCGAGCTCGACAGCGTGCAGACGGTGTGATCTCT 280422
55ArgLeuAlaProAsnPheA 61
280423 TACCGCGGATAGCGCCACCGATTGGCGCGCGCTCGCTTCGGTCAGC 280472
61 rgtPralleProAsnArgGlnMetAsnAspGlyLeuGly..... 75
280473 GTGGCCCGCGGACCGCGGGATACAGCAGGATGCTTCGACGGCTCCGA 280522
76AspGlyAs 78
280523 GCTAACAGCACGCTCTGGAAGCAGCGCTCCGGAGACCGCTCGGGCA 280572
78 pAspMetGluMetPheMetGluLeuMetArgGluLeuArgGlyLeuA 95
280573 TGACGTGGCGAGCTACAGCGCGCATATTCGCGGACGTTTCGACGGCTGTTG 280622


```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 663..3164
; US-08-396-001-3

alignment_scores:
  Quality: 79.00      Length: 92
  Ratio: 1.362      Gaps: 5
  Percent Similarity: 63.043      Percent Identity: 26.087

alignment_block:
  US-09-327-750D-12 x US-08-396-001-3 ..
  Align seg 1/1 to: US-08-396-001-3 from: 1 to: 3455

17 Asnglyglucluaspaqprvaiglyglygluyluaglnproal 33
  17 AACACGAAATAGAACCC.....CATACAAAAA 2482
  33 aGlyAsnAsnAsnAsnAsnAsnHIsAsnHIsAsnHIsArGA 50
  2483 TAGTCATATCATATCATATCATATCATATCATATCATAT 2532
  50 rGlyGlnAlaArGArGLeuAlaProAsnPhaArgTrpAlaIlePro... 65
  2533 ATACATATATATATATATATATATATATATATATATATAT 2582
  66 .....AsnArgGlnMetAsnAsnGlyLeuGlyAspGlyAspAs 79
  2583 ATGCTTACCATAGAGAGATACACCTCTCTTACCAATAT 2624
  79 pMetGluMetPhaMetGluGluMetArgGluIleArGArGlyLeuArG 95
  2625 .....TTCTCAACCAATATG.....CACAGATCAGA 2652
  96 GluLeuGlnLeuArgAsnCysLeu 103
  2653 AAATTCACCTCTCCGACCAATTA 2676

seq_name: /cgn2_6/pdata/2/lna/68_COMB.seq:US-09-323-433A-3

seq_documentation_block:
; Sequence 3, Application US/09323433A
; Patent No. 6218512
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: DETERMINING CELLULAR SENSICENCE IN
; FILE REFERENCE: 0050.1491-003
; CURRENT APPLICATION NUMBER: US/09/323,433A
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 1: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3: FASTSEQ for Windows Version 3.0
; LENGTH: 3455
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (663)...(3164)
```

```
; OTHER INFORMATION: UTH4
; US-09-323-433A-3

alignment_scores:
  Quality: 79.00      Length: 92
  Ratio: 1.362      Gaps: 5
  Percent Similarity: 63.043      Percent Identity: 26.087

alignment_block:
  US-09-327-750D-12 x US-09-323-433A-3 ..
  Align seg 1/1 to: US-09-323-433A-3 from: 1 to: 3455

17 Asnglyglucluaspaqprvaiglyglygluyluaglnproal 33
  17 AACACGAAATAGAACCC.....CATACAAAAA 2482
  33 aGlyAsnAsnAsnAsnAsnAsnHIsAsnHIsAsnHIsArGA 50
  2483 TAGTCATATCATATCATATCATATCATATCATATCATAT 2532
  50 rGlyGlnAlaArGArGLeuAlaProAsnPhaArgTrpAlaIlePro... 65
  2533 ATACATATATATATATATATATATATATATATATATATAT 2582
  66 .....AsnArgGlnMetAsnAsnGlyLeuGlyAspGlyAspAs 79
  2583 ATGCTTACCATAGAGAGATACACCTCTCTTACCAATAT 2624
  79 pMetGluMetPhaMetGluGluMetArgGluIleArGArGlyLeuArG 95
  2625 .....TTCTCAACCAATATG.....CACAGATCAGA 2652
  96 GluLeuGlnLeuArgAsnCysLeu 103
  2653 AAATTCACCTCTCCGACCAATTA 2676

seq_name: /cgn2_6/pdata/2/lna/6A_COMB.seq:US-08-155-888-1

seq_documentation_block:
; Sequence 1, Application US/08155888
; Patent No. 6068333
; GENERAL INFORMATION:
; APPLICANT: Hoffmann, Stephen L.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical Res. & Dev. Cnd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,888
; FILING DATE:
; PRIORITY DATE:
; PRIORITY APPLICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
```

```
561 .....CCCCAGCGCCAGCCT..... 575
80 exGluMetPheMetGluMetArgGluLeuLeuArgLysLeuArgGlu 96
   |||||.....
576 .....CTACCGCTCAAGTACTCGCGCCGACGACGCGCCCA 614
97 LeuGlnLeuArgAspCysLeuArgLeuMetGlyGlyGlyLeuSerAsnH1 113
   |||||.....
615 GTACAGCTGGATCAT...CAACGGCATCGAGTGGCCATCGCCACAACA 661
113 sHisAspHisHisAspGlu 119
   |||||.....
662 TGGAGTCATCACCATTGAG 680

seq_name: /cgn2_6/podata/2/ina/5B_COMB.seq:US-08-387-942C-18
seq_documentation_block:
; Sequence 18, Application US/08387942C
; Patent No. 5939289
; GENERAL INFORMATION:
; APPLICANT: ERTESVAG, HELGA
; APPLICANT: VALLA, SVEIN
; APPLICANT: SKJAK-BRAEK, GUDMUND
; APPLICANT: LARSEN, BJORN
; TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
; OF SEQUENCES 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000, 1001, 1002, 1003, 1004, 1005, 1006, 1007, 1008, 1009, 1010, 1011, 1012, 1013, 1014, 1015, 1016, 1017, 1018, 1019, 1020, 1021, 1022, 1023, 1024, 1025, 1026, 1027, 1028, 1029, 1030, 1031, 1032, 1033, 1034, 1035, 1036, 1037, 1038, 1039, 1040, 1041, 1042, 1043, 1044, 1045, 1046, 1047, 1048, 1049, 1050, 1051, 1052, 1053, 1054, 1055, 1056, 1057, 1058, 1059, 1060, 1061, 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1070, 1071, 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1081, 1082, 1083, 1084, 1085, 1086, 1087, 1088, 1089, 1090, 1091, 1092, 1093, 1094, 1095, 1096, 1097, 1098, 1099, 1100, 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109, 1110, 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120, 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129, 1130, 1131, 1132, 1133, 1134, 1135, 1136, 1137, 1138, 1139, 1140, 1141, 1142, 1143, 1144, 1145, 1146, 1147, 1148, 1149, 1150, 1151, 1152, 1153, 1154, 1155, 1156, 1157, 1158, 1159, 1160, 1161, 1162, 1163, 1164, 1165, 1166, 1167, 1168, 1169, 1170, 1171, 1172, 1173, 1174, 1175, 1176, 1177, 1178, 1179, 1180, 1181, 1182, 1183, 1184, 1185, 1186, 1187, 1188, 1189, 1190, 1191, 1192, 1193, 1194, 1195, 1196, 1197, 1198, 1199, 1200, 1201, 1202, 1203, 1204, 1205, 1206, 1207, 1208, 1209, 1210, 1211, 1212, 1213, 1214, 1215, 1216, 1217, 1218, 1219, 1220, 1221, 1222, 1223, 1224, 1225, 1226, 1227, 1228, 1229, 1230, 1231, 1232, 1233, 1234, 1235, 1236, 1237, 1238, 1239, 1240, 1241, 1242, 1243, 1244, 1245, 1246, 1247, 1248, 1249, 1250, 1251, 1252, 1253, 1254, 1255, 1256, 1257, 1258, 1259, 1260, 1261, 1262, 1263, 1264, 1265, 1266, 1267, 1268, 1269, 1270, 1271, 1272, 1273, 1274, 1275, 1276, 1277, 1278, 1279, 1280, 1281, 1282, 1283, 1284, 1285, 1286, 1287, 1288, 1289, 1290, 1291, 1292, 1293, 1294, 1295, 1296, 1297, 1298, 1299, 1300, 1301, 1302, 1303, 1304, 1305, 1306, 1307, 1308, 1309, 1310, 1311, 1312, 1313, 1314, 1315, 1316, 1317, 1318, 1319, 1320, 1321, 1322, 1323, 1324, 1325, 1326, 1327, 1328, 1329, 1330, 1331, 1332, 1333, 1334, 1335, 1336, 1337, 1338, 1339, 1340, 1341, 1342, 1343, 1344, 1345, 1346, 1347, 1348, 1349, 1350, 1351, 1352, 1353, 1354, 1355, 1356, 1357, 1358, 1359, 1360, 1361, 1362, 1363, 1364, 1365, 1366, 1367, 1368, 1369, 1370, 1371, 1372, 1373, 1374, 1375, 1376, 1377, 1378, 1379, 1380, 1381, 1382, 1383, 1384, 1385, 1386, 1387, 1388, 1389, 1390, 1391, 1392, 1393, 1394, 1395, 1396, 1397, 1398, 1399, 1400, 1401, 1402, 1403, 1404, 1405, 1406, 1407, 1408, 1409, 1410, 1411, 1412, 1413, 1414, 1415, 1416, 1417, 1418, 1419, 1420, 1421, 1422, 1423, 1424, 1425, 1426, 1427, 1428, 1429, 1430, 1431, 1432, 1433, 1434, 1435, 1436, 1437, 1438, 1439, 1440, 1441, 1442, 1443, 1444, 1445, 1446, 1447, 1448, 1449, 1450, 1451, 1452, 1453, 1454, 1455, 1456, 1457, 1458, 1459, 1460, 1461, 1462, 1463, 1464, 1465, 1466, 1467, 1468, 1469, 1470, 1471, 1472, 1473, 1474, 1475, 1476, 1477, 1478, 1479, 1480, 1481, 1482, 1483, 1484, 1485, 1486, 1487, 1488, 1489, 1490, 1491, 1492, 1493, 1494, 1495, 1496, 1497, 1498, 1499, 1500, 1501, 1502, 1503, 1504, 1505, 1506, 1507, 1508, 1509, 1510, 1511, 1512, 1513, 1514, 1515, 1516, 1517, 1518, 1519, 1520, 1521, 1522, 1523, 1524, 1525, 1526, 1527, 1528, 1529, 1530, 1531, 1532, 1533, 1534, 1535, 1536, 1537, 1538, 1539, 1540, 1541, 1542, 1543, 1544, 1545, 1546, 1547, 1548, 1549, 1550, 1551, 1552, 1553, 1554, 1555, 1556, 1557, 1558, 1559, 1560, 1561, 1562, 1563, 1564, 1565, 1566, 1567, 1568, 1569, 1570, 1571, 1572, 1573, 1574, 1575, 1576, 1577, 1578, 1579, 1580, 1581, 1582, 1583, 1584, 1585, 1586, 1587, 1588, 1589, 1590, 1591, 1592, 1593, 1594, 1595, 1596, 1597, 1598, 1599, 1600, 1601, 1602, 1603, 1604, 1605, 1606, 1607, 1608, 1609, 1610, 1611, 1612, 1613, 1614, 1615, 1616, 1617, 1618, 1619, 1620, 1621, 1622, 1623, 1624, 1625, 1626, 1627, 1628, 1629, 1630, 1631, 1632, 1633, 1634, 1635, 1636, 1637, 1638, 1639, 1640, 1641, 1642, 1643, 1644, 1645, 1646, 1647, 1648, 1649, 1650, 1651, 1652, 1653, 1654, 1655, 1656, 1657, 1658, 1659, 1660, 1661, 1662, 1663, 1664, 1665, 1666, 1667, 1668, 1669, 1670, 1671, 1672, 1673, 1674, 1675, 1676, 1677, 1678, 1679, 1680, 1681, 1682, 1683, 1684, 1685, 1686, 1687, 1688, 1689, 1690, 1691, 1692, 1693, 1694, 1695, 1696, 1697, 1698, 1699, 1700, 1701, 1702, 1703, 1704, 1705, 1706, 1707, 1708, 1709, 1710, 1711, 1712, 1713, 1714, 1715, 1716, 1717, 1718, 1719, 1720, 1721, 1722, 1723, 1724, 1725, 1726, 1727, 1728, 1729, 1730, 1731, 1732, 1733, 1734, 1735, 1736, 1737, 1738, 1739, 1740, 1741, 1742, 1743, 1744, 1745, 1746, 1747, 1748, 1749, 1750, 1751, 1752, 1753, 1754, 1755, 1756, 1757, 1758, 1759, 1760, 1761, 1762, 1763, 1764, 1765, 1766, 1767, 1768, 1769, 1770, 1771, 1772, 1773, 1774, 1775, 1776, 1777, 1778, 1779, 1780, 1781, 1782, 1783, 1784, 1785, 1786, 1787, 1788, 1789, 1790, 1791, 1792, 1793, 1794, 1795, 1796, 1797, 1798, 1799, 1800, 1801, 1802, 1803, 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813, 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823, 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833, 1834, 1835, 1836, 1837, 1838, 1839, 1840, 1841, 1842, 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850, 1851, 1852, 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862, 1863, 1864, 1865, 1866, 1867, 1868, 1869, 1870, 1871, 1872, 1873, 1874, 1875, 1876, 1877, 1878, 1879, 1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889, 1890, 1891, 1892, 1893, 1894, 1895, 1896, 1897, 1898, 1899, 1900, 1901, 1902, 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1910, 1911, 1912, 1913, 1914, 1915, 1916, 1917, 1918, 1919, 1920, 1921, 1922, 1923, 1924, 1925, 1926, 1927, 1928, 1929, 1930, 1931, 1932, 1933, 1934, 1935, 1936, 1937, 1938, 1939, 1940, 1941, 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 
```


553 TATATGCGCTTCMAACTCTCGACCTGACACCATTTATGAC..... 594

118 pGluPheCysLeuMetPro 124

|||||.....

595TTTTGCTCATACCT 609

seq_name: /SIDS2/jcgdata/geneseq/geneseq/NA2000.DAT:AAH87147

seq_documentation_block:

ID AAH87147 standard: DNA: 187 BP.

XX AAH87147:

XX 08-JAN-2001 (first entry)

DE Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:71.

KW Rat: phenobarbital; carcinogenesis marker; carcinogenesis; detection;

KW Identification; carcinogenic; probe; primer; ds.

OS Rattus norvegicus.

XX WO200044902-A2.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US00503.

XX 29-JAN-1999; 99US-0118078.

XX (SEAR) SEARLE & CO G D.

XX Bunch RT, Curtis SW, Rodi CP, Morris DL;

XX WPI: 2000-505977/45.

XX New nucleic acid encoding a carcinogenic biomarker, induced by

PT phenobarbital treatment of rat hepatocytes, useful for identifying

PT carcinogenic compounds.

PS Claim 1; Page 73; 240pp; English.

CC AAH87080 to AAH87656 represent nucleic acid sequences (N1) encoding a

CC carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by

CC treating rat hepatocytes with phenobarbital. The nucleic acids are

CC useful for identifying carcinogenic compounds. The nucleic acids are

CC used for detecting and/or primers for detecting or inducing

CC carcinogenesis, respectively.

SQ Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other;

alignment_scores:

Quality: 193.00 Length: 35

Ratio: 5.514 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-12 x AAH87147

Align seg 1/1 to: AAH87147 from: 1 to: 187

90 ILeArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLe 106

|||||.....

2 ATCCGGAGAAAGCTTAGGGAGCTGCAGTTGAGAAATTCCTCGTATCT 51

|||||.....

106 uMetClyGluLeuSerAsnHisHisAspHisAspGluPheCysLeuM 123

|||||.....

52 TATCGGGAGCTCTCTATATCCACCCACCATCGCATGATTCCTGCTTA 101

|||||.....

102 TGCT 106

seq_name: /SIDS2/jcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard: cDNA: 792 BP.

XX AAH45143:

XX 07-SEP-2001 (first entry)

XX Human brain expressed X-linked protein, hhex, coding sequence.

KW Human: brain expressed X-linked protein; cytostatic; auditory; nontropic;

KW hhex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;

KW X-chromosome-binding mental retardation; lissencephalous disease; ss.

OS Homo sapiens.

XX WO200140286-A1.

XX 07-JUN-2001.

XX 27-NOV-2000; 2000WO-CN00502.

XX 30-NOV-1999; 99CN-0124179.

XX (BIOR-) BIORAD GENE DEV LTD SHANGHAI.

XX Mao Y, Xie Y;

XX WPI: 2001-397944/42.

XX P-PSDB: AAB9224.

XX Isolated human brain-expressed X-linked polypeptide used to diagnose

PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor,

PT deafness and X-chromosome-binding mental retardation.

PS Claim 5; Page 22; 30pp; Chinese.

XX The present sequence is the coding sequence for a human brain-expressed

CC X-linked protein (hhex). hhex and its coding sequence are useful in the

CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,

CC tumours, deafness, X-chromosome-binding mental retardation and

CC lissencephalous disease. hhex is also useful for identifying

CC genes encoding or the coding sequence of the gene for fingerprint identification.

CC Sequencing sequences can be used as primers or probes, or in producing

CC gene chips or microarrays.

SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:

Quality: 184.00 Length: 131

Ratio: 2.452 Gaps: 5

Percent Similarity: 57.252 Percent Identity: 36.642

alignment_block:

US-09-327-750D-12 x AAH45143

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17

|||||.....

212 ATGGAAATGCGCAACCAAGAAATGAGAAAGGAGCAGTTCCTAATPA 261

|||||.....

17 nGlyGluGlu...AspArgProValGlyGlyGlyGlyHisGlnProk 33

|||||.....

262 AGGGAGCCCTTCGCCCTCCCTTTGGATGCTGGTGAATCTGTGTGCTTA 311

|||||.....

33 LacIlyAnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsn 49

|||||.....

312 GAGGAAT.....CGTAGG 325

XX	AF59611 standard; cDNA; 898 bp.
XX	
XX	AF59611;
XX	
XX	24-APR-2001 (first entry)
XX	
DE	Human cell cycle and proliferation protein CCYP8-22 cDNA, SEQ ID NO.76.
XX	
XX	Cell cycle and proliferation protein; CCYP8; human; agonist;
KW	antagonist; gene therapy; detection; gene therapy;
KW	developmental disorder; cell signalling disorder;
KW	developmental disorder; cell signalling disorder;
KW	cell proliferative disorder; cancer; tumour; anaemia;
KW	arteriosclerosis; asthma; allergy; diabetes mellitus;
KW	menstrual cycle disorder; bacterial infection; ss.
XX	
XX	Homo sapiens.
XX	
PN	WO200107471-A2.
XX	
XX	01-FEB-2001.
PD	
XX	
XX	21-JUL-2000; 2000WO-US19948.
XX	

PK 21-JUL-1999: 99US-0153073.
PK 08-SEP-1999: 99US-0153129.

XX
PA PA .(INCY-) INCYTE GENOMICS INC.
XX
PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;

XX
DR
WPI: 2001-112727/12.

Human cell cycle and proliferation proteins and polynucleotides are

PT used to treat, diagnose and prevent immune, developmental and cell signaling disorders and cell proliferative disorders including cancer -

XX
PS
Claim 5: Page 181-182: 205pp: English.

CC
XX
XXXXXXXXXXXXX ADE500-AD500? respondent CNA's encoding 54 human

CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.

CC associated with decreased expression of functional CCYPR, while CCYPR

CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies

radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect

CC compounds e.g., antibodies, oligonucleotides and proteins (receptors) that specifically bind to CCYPR, and in drug screening methods to

CC identity compounds that modulate the activity of CCPR. CCPR nucleotides can be used to generate transgenic animal models of human

CC disease, and can be used in gene therapy in target cells with genetic abnormalities with respect to the expression of CCYPB for the

CC treatment or prevention of a disorder associated with CCyPR.
CC diseases which can be diagnosed, treated and prevented using CCyPR

CC proteins, nucleic acids, agonists or antagonists include immune, developmental and cell signalling disorders and cell proliferative

disorders including cancer. Specific examples of these disorders

CC diabetes mellitus, disorders of the menstrual cycle and infections

XX
XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

alignment_scores:

Quality:	171.50
Ratio:	2.287
Length:	132
Gaps:	6

Percent Similarity: 56.818 Percent Identity: 36.364

obtained the full 5' UTR is rarely included, 5' ESTs are derived from CC
CC DNA with intact 5' ends and can therefore be used to obtain full length CC
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, CC
CC gene therapy and chromosome mapping procedures. They are used to obtain CC
CC upstream regulatory sequences and to design expression and secretion CC
CC vectors.

```

alignment_scores:      Length: 126
                      Quality: 162.50
                      Ratio: 2.355
                      Gaps: 4
Percent Similarity: 54.762 Percent Identity: 35.714

alignment_block:
US-09-327-750D-12 x AAC03880 ..

Align seq 1/1 to: AAC03880 from: 1 to: 662

3 AsnValHisGlnGluAsn.....GluGluMetGluGlnProLeuG1 16
252 ANTGAGTGGAGGAAAGAAAGTGTGTAAGAAAGTGTGTAAGTACAGCTACGCTAA 301
16 nAsnGlyGluGlu...AspATgProValGlyGlyGlyGluHisGlnP 32
302 TAAAGGGGAGCCCTTGCGCTACCTTTGTGATGTTGTGTGAATCTGTGTC 351
32 roAlaGlyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisH 48
352 CTAGAGGAAT.....CGT 365
49 ArgArgGlyGluAlaArgArgLeuAlaProAsnPheArgTTPAlaLepr 65
366 AGGCGCTTCGGGTGTGGTACCCATCTCTCATGTATGATGGATATGT 415
65 oAsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAspMetGluM 82
416 GCATAGGCTTGGAGACCAAGCAGGACAGGATGAGAGAGAGATATGAAA 465
82 ePheMetGlyGluMetArgGluGluLeuArgArgGlyLeuArgGluLeuGln 98
466 GGATTTGGGAGGAGGTGTGACACACTGTATGTAAGAAAGCTGTAGGGAAGACAG 515
99 LeuArgAsnCysLeuArgTLeuLeuMetGlyGlyLeuLeuSerAsnHisHisAs 115
516 TTGTGATCATGCTGTGGCGGCGAGTGCACGACTGC 124.....CCCCCTCACCTTCA 562
115 pHisHisAspGluPheCysLeuMetpro 124
563 GCATCATGATGAGTATTTTGCNNNNATGCC 590

```


Strd	Orig	zScore	Len	Documentation	...
+	688.00	1064.22	5.3e-51	700	AF187066 Mus musculus p75NTR-se
+	688.00	1064.22	5.3e-51	700	AF187066 Mus musculus p75NTR-se
+	688.00	1062.65	6.5e-51	854	AF097440 Mus musculus brain exp
+	643.00	997.06	2.9e-47	519	AF097440 Mus musculus brain exp
+	643.00	997.06	2.9e-47	519	AF097440 Mus musculus brain exp
+	576.50	890.99	2.4e-41	793	BC003190 Homo sapiens, p75NTR-4
+	576.50	890.99	2.4e-41	793	BC003190 Homo sapiens, p75NTR-4
+	576.50	890.07	2.4e-41	891	K18188 Human unknown protein fr
+	576.50	890.07	2.4e-41	891	K18188 Human unknown protein fr
+	411.00	640.31	2.6e-27	421	G72708 HMC h93c14594 93191003
+	411.00	640.31	2.6e-27	421	G72708 HMC h93c14594 93191003
+	382.00	404.88	2.8e-14	45678	G72708 HMC h93c14594 93191003
+	382.00	404.88	2.8e-14	45678	G72708 HMC h93c14594 93191003
+	364.50	405.55	2.6e-14	1329	AF000959 Homo sapiens cDNA FL1
+	364.50	405.55	2.6e-14	1329	AF000959 Homo sapiens cDNA FL1
+	364.50	404.73	2.9e-14	1364	AF100231 Sequence 32 from Paten
+	364.50	404.73	2.9e-14	1364	AF100231 Sequence 32 from Paten
+	440.50	333.97	1.5e-12	69648	Z599755 Human DNA sequence fr
+	440.50	333.97	1.5e-12	69648	Z599755 Human DNA sequence fr
+	440.50	333.97	1.5e-12	98274	G53524 human SRS SHC3-74/09, be
+	440.50	333.97	1.5e-12	98274	G53524 human SRS SHC3-74/09, be
+	191.00	297.12	2.4e-02	612	AF201189 Homo sapiens osteonin
+	191.00	297.12	2.4e-02	612	AF201189 Homo sapiens osteonin
+	191.00	295.12	3.7e-08	838	AF201189 Homo sapiens osteonin
+	191.00	295.12	3.7e-08	838	AF201189 Homo sapiens osteonin
+	184.00	265.47	1.4e-07	35714	AL008708 Human DNA sequence fr
+	184.00	265.47	1.4e-07	35714	AL008708 Human DNA sequence fr
+	172.00	266.19	1.5e-06	785	AF237783 Homo sapiens brain-exp
+	172.00	266.19	1.5e-06	785	AF237783 Homo sapiens brain-exp
+	171.50	265.84	1.6e-06	744	AF251053 Homo sapiens x-linked
+	171.50	265.84	1.6e-06	744	AF251053 Homo sapiens x-linked
+	171.50	264.35	1.9e-06	898	AF078372 Sequence 76 from Paten
+	171.50	264.35	1.9e-06	898	AF078372 Sequence 76 from Paten
+	171.50	234.34	8.9e-05	40564	AF133334 Human DNA sequence fr
+	171.50	234.34	8.9e-05	40564	AF133334 Human DNA sequence fr
+	171.50	234.25	9.0e-05	41029	AL022169 Homo sapiens chromos
+	171.50	234.25	9.0e-05	41029	AL022169 Homo sapiens chromos
+	156.00	240.98	3.8e-05	835	AF097438 Homo sapiens brain exp
+	156.00	240.98	3.8e-05	835	AF097438 Homo sapiens brain exp
+	156.00	240.95	3.8e-05	838	AF051347 Mus musculus REX-3 mRN
+	156.00	240.95	3.8e-05	838	AF051347 Mus musculus REX-3 mRN
+	156.00	233.11	0.0001	2269	AF097437 Mus musculus Bex1 pr
+	156.00	233.11	0.0001	2269	AF097437 Mus musculus Bex1 pr
+	152.50	204.36	0.0042	43952	Z598246 Homo sapiens chrom
+	152.50	204.36	0.0042	43952	Z598246 Homo sapiens chrom
+	138.00	172.06	0.2626	14577	AF002162 Homo sapiens clone
+	138.00	172.06	0.2626	14577	AF002162 Homo sapiens clone
+	138.00	170.63	0.1033	18559	AF002162 Homo sapiens clone
+	138.00	170.63	0.1033	18559	AF002162 Homo sapiens clone
+	122.50	193.21	0.0174	504	G24644 human SRS M1-135 sequ
+	122.50	193.21	0.0174	504	G24644 human SRS M1-135 sequ
+	112.00	166.40	0.5425	19393	AF181716 Cynomolgus Epstein-Ba
+	112.00	166.40	0.5425	19393	AF181716 Cynomolgus Epstein-Ba
+	106.50	132.68	95.54	109891	Z589971 Human DNA sequence fr
+	106.50	132.68	95.54	109891	Z589971 Human DNA sequence fr
+	106.12	131.35	175.26	321058</	

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 793)
Direct Submission
Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2950,
USA
MGC project URL: <http://mgc.nci.nih.gov>
Contact: GAGB help desk
Email: gcgbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Gary Garbutt, Guo
Sheng, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
Ness, Pawan Pandoh, Anna-Liia Rabbu, Parvaneh Saedi, Jacqueline
Schain, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E Consortium/LLNL at: http://image.llnl.gov
Series: IRAL plate: 6 Row: k Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7657043.

FEATURES
    source
        1..793
            /organism="Homo sapiens"
            /db_xref="LocusID:27018"
            /db_xref="taxon:9606"
            /clone="MGC-802 IMAGE:3357965"
            /tissue_type="Eye, retinoblastoma"
            /clone.lib="IRAL-MGC-16"
            /db_xref="GeneID:101083"
            /db_xref="vector: pORF7"
            /codon_start=1
            /product="p53NTR-associated cell death executor; ovarian
            granulosa cell protein (13kD)"
            /protein_id="AAH03190.1"
            /db_xref="GI:13112031"
            /translation="MANTHQENEMEDPQMONGEDRPLGGSGHGPAGNRGRGAARLLA
            PNRWATPNQINQNGDGDGDEIFEMREIRKLQRLNCLRLIMGELSGSLHIA
            HHDFECLMP"
BASE COUNT      225 a   205 g   193 t
ORIGIN

alignment_scores
    Quality: 576.50      Length: 124
    Ratio: 5.194         Gaps: 1
    Percent Similarity: 89.516   Percent Identity: 84.677

alignment_block:
    US-09-327-7500-12 x BC003190 ..
Align seq 1/1 to: BC003190 from: 1 to: 793

1 MetAlaenValHicGInGluangGluGluWecGluGInProLeuGlnAs 17
|||||
196 ATGGCAATATTCAGGAGAGACGAGATGATGGAGGCGCTATGCGAGAA 245
|||||

17 nclYGluAspArgProValGlyGlyGlyGluGlyHISgInProIaG 34
|||||
246 TGGAGAGAGACGCCCTTTTGGAGAGAGGTTGAAGGCCACCAACCGCTGCA 295
|||||

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..891
/gene="NAGE"
312..640g
/contig="NAGE"
/contig="involved in the common neurotrophin receptor
p75NTR-mediated signal transduction"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75129.1"
/db_xref="GI:8452894"
/translation="MANIHSENEEMQPMQNEEDRPLGGEGHQPAGNRRQARLA
PNFRNAIPNRIINDGGGDDHDEIFMEHREIRKLHQLANCLRLMGLSELNHH
HIDEFLMP"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN

alignment_scores:
  Quality: 576.50      Length: 124
  Ratio: 5.194         Gaps: 1
Percent Similarity: 89.516 Percent Identity: 84.677

alignment_block:
US-09-327-750D-12 x AF187064 ..
Align seg 1/1 to: AF187064 from: 1 to: 891

1 MetAlaAanValHisGluGluAargGluMetGluGluProLeuGlnAs 17
312 ATGGCAATATTCCACGAAAGAGAGAGATGGAGCAGCTATGCGAA 361
17 nCylGluAAspArgProValGlyGlyGlyGlyHisGlnProAlaG 34
362 TGGAGGAGGAGACCCCTTTGGGAGGAGGTGAAGCCACGACCTGCAG 411
34 LyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArg 50
412 GAAT .....CCACGG 422
51 GylGlnAlaArgLeuAlaProAnPheArgTrpAlaIleProAnAr 67
423 GGACAGCTCGGACCTGCCCTATTTTCAGTGGCCATACCCATAG 472
67 GlnMetAsnAspGlyLeuGlyGlyGlyGlyAspGlyAspMetGluMetPheM 84
473 GCAGATCAATGATGGATGGGTGGAGATGGAGATGATATGGAATATCA 522
84 eGluGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArg 100
523 TGGAGGAGATGAGAAATACAGAGAAACTTACGAGGCTGAGTTGAGG 572
101 AnCysLeuArgIleLeuMetGlyGlyLeuSerAsnHisAspHisI 117
573 AATTGCTGCTATCTTATGGGGAGCTCTCTATACCATGACCATCA 622
117 sAspGluPheCysLeuMetPro 124
623 TGATGAATTTTGCCTTATGCT 644

seq_name: gb_sts:G72708
```

```
seq_documentation_block:
LOCUS G72708 421 bp DNA STS 08-AUG-2001
DEFINITION MARC 4953:4954:99193031:1 SCF - porcine spleen STS
ACCESSION G72708.1
VERSION G72708.1 GI:15146738
KEYWORDS STS.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 421)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L. and
Keefe,J.W.
TITLE
Single nucleotide polymorphism (SNP) discovery in expressed porcine
DNA
JOURNAL
Unpublished (2001)
COMMENT
Contact: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278
Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: GCACATGGGTGAACTACT
Primer B: CCACGACCAATAGACG
STS size: 300
PCR Protocol:
Denature: 95 degrees for 15 minutes
Anneal: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45
Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmoles
dNTPs: each 88 uM
Taq Polymerase: 0.25 units (Qiagen HotStar)
Buffer:
Commercially supplied Qiagen HotStar buffer
The STS is derived from PCR amplicons generated from genomic DNA,
sequenced from each end using the amplification primers. The
sequence does not necessarily represent the entire amplicon.
Sequence derived from PolyPhred was trimmed from each end of each
unique contig until five consecutive bases exceeded a quality score
threshold of 20, and the next 10 bases averaged a quality score of
20 or greater. Amplicon size was estimated by agarose gel
electrophoresis.
Location/Qualifiers
1..421=Sun scrofa"
/organism="Sus scrofa"
/strain="white composite, duroc, meishan, minzhu,
fengling, crossbred"
/db_xref="taxon:9823"
/sex="male and female"
/clone_lib="SCF - porcine spleen"
/dev_stage="adult"
/note="Organ: spleen"
<1..421
BASE COUNT 170 a 101 c 86 g 102 t 2 others
ORIGIN

alignment_scores:
  Quality: 411.00      Length: 77
  Ratio: 5.338         Gaps: 0
Percent Similarity: 100.000 Percent Identity: 96.104

alignment_block:
US-09-327-750D-12 x G72708/rev ..
Align seg 1/1 to reverse of: G72708 from: 1 to: 421

48 HisArgGlyGlyGlnAlaArgLeuAlaProAnPheArgTrpAlaI 64
421 ATAGACCGGACAGCTGCCCTATTTTCAGTGGCCATACCCATG 372
64 eProAnArgGlnMetAsnAspGlyLeuGlyGlyGlyAspAspMetG 81
371 ACCCAATAGCAGATCAATGATGGATGGGTGGAGATGGAGATGAT 322
```

seq_name: gb_pr:AK000959

seq_documentation_block: 1229 bp mRNA PRI 22-FEB-2000

LOCUS AK000959

DEFINITION Homo sapiens cDNA FLJ10097 fis, clone HEMBA1002458, weakly similar to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.

ACCESSION AK000959

VERSION 1

KEYWORDS GI:7021945

SOURCE Homo sapiens: fetus (full insert sequence).

ORGANISM Homo sapiens

REFERENCE 1. Ota, T., Hayashi, K., Sugiyama, T., Otauki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Kometani, F., Haru, R., Takeuchi, K., Arica, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuhara, Y. and Oshima, A. NEDO human cDNA sequencing project

db-esc2-BI281721	538.00	1061.48	5.86-50	662	BI281721	UI-R-CT09-cav.f.
db-esc2-BI189828	621.50	1048.73	3.08-49	810	BF169828	6017375392FI NC1
db-esc2-AZ14909	618.00	1046.43	4.00-49	532	AA14909	um7605.r1 Strat
db-esc1-AI183075	615.00	1042.69	6.56-49	455	AI183075	ub3960.r1 Soare
db-esc1-AI157489	611.00	1037.80	1.26-48	366	AI157489	es56605.r1 Soare

Accession name: gb_ests:W45041

Accession name	Accession number	Accession type	Accession date	Accession status	Accession description	Accession length	Accession size	Accession date	Accession status
gb_ests:W45041	492 bp	mRNA	EST	23-MAY-1996	Sequence documentation block: LOCUS M60401 DEFINITION mc82h02.r1 Soares mouse embryo NM013.5 14.5 Mus musculus CDNA clone IMAGE355059.57, similar to gb:M31818 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN); / gb:NM013.5 OVARIAN	492	492	23-MAY-1996	EST

TITLE	JOURNAL	COMMENT
1. (bases 1 to 492)		
Maria, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Giesel, S., Kuchaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Septemo, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Taborson, T.		
The NIH-HMI Mouse EST Project		
Unpublished (1996)		
Contact: Maria M/Mouse EST Project		
WashU-HMI Mouse EST Project		
Washington University School of Medicine		
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
Tel: 314 286 1800		

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouse@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouse@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:226859
 Seq primer: ETPRimer
 High quality sequence stop: 489.
 Location/Qualifiers
 1..492
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /taxon="10090"
 /clone_accession="MIM235059"
 /clone_lib="Scars mouse embryo NDBE13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /vector="pET7D3-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not; Site_2: Eco RI; lat strand cDNA
 was primed with a Not I - oligo(dT) primer [5]
 TGTTACCACTGATGGGAGGCGCGGGAATTTTTTTTTTTTTTTTTTTTTT
 T 3', on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 P. Alivisatos, and Ronan O'Keefe, with the aid
 of E. R. Skaar, and the National Cancer Institute].
 Cloned into the Not I and Eco RI sites of the modified
 p7773 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M. Fatima Bonaldo.
 133 c 125 c 101 t

```

alignment_scores:
  Quality: 688.00
  Ratio: 5.348
  Percent Similarity: 100.000
  Length: 124
  Gaps: 0
  Percent Identity: 98.387
alignment block:

```

alignment_block:


```

VERSION BE334866.1 GI:9208642
KEYWORDS EST
SOURCE house mouse
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 342)
COMMENT NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
MGI:1069792
Seq primer: -40RP from Gibco
High quality sequence stop: 465.
FEATURES
source
1..342
Location/Qualifiers
tag="mus musculus"
/clone="IMAGE:3325628"
/clone_lib="Soares mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - OligodT primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia); digested with Not
I; cloned into the pT73D-Pac vector. Library is small-scale
modified pT73D vector. Library is small-scale library
was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 156 a 141 c 164 g 81 t
ORIGIN

alignment_scores
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block
US-09-327-750D-12 x BE334866
Align seg 1/1 to: BE334866 from: 1 to: 542
1 MetAlaAnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
166 ATGGCCATATCTCCACAGGAAACAGAGAGATGAGCAGCCCTGCAGAA 215
17 nGlyGluAspArgProValGlyGlyGlyGlyHisGlnProAlaG 34
216 TGGACGAGAGACCCCTGTGGAGAGAGTGGGACCACCCCTCTGTG 265
34 LyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArg 50
266 CAAACACACACACACACACACACACACACACACACACACACAG 315
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
316 GGCCAGGCTGCCGACCTCCCTTAACCTCCGATGGCCATCCCAACAG 365
67 gGlnMetAlaAspGlyLeuGlyGlyAspGlyAspAspMetGluMetPheH 84
366 CCGATGATGACGGGTGGCTGGAGATGGGATGATATGGAAATGTTC 415
84 etGluGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArg 100
416 TGGAGGAGATGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 465

```

```

101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
117 aAspGluPheCysLeuMetPro 124
465 AATTGTCTACGATTCTTATGGGGAGCTGTCTTAACCCACCGATCACA 515
117 aAspGluPheCysLeuMetPro 124
516 TGATGATTTCTGCTTATGCTCTTATGCTCTTATGCTCTTATGCTCTT 557
seq_name: gb_est1:AA215070
seq_documentation_block:
LOCUS AA215070 557 bp mRNA EST 03-FEB-1997
DEFINITION musculus cDNA clone IMAGE:551487.5' similar to gb:M38188 OVARIAN
musculus cDNA clone IMAGE:551487.5' similar to gb:M38188 OVARIAN
GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION AA215070
VERSION AA215070.1 GI:1814831
KEYWORDS EST
SOURCE house mouse
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 557)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Marta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Giesel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, P., Underwood, K., Moore, B.,
Theising, B., Wyllie, T., Lannon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washington University School of Medicine
660 South Euclid Avenue, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
MGI:397335
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 447.
FEATURES
source
1..557
Location/Qualifiers
/organism="Mus musculus"
/clone="IMAGE:551487"
/clone_lib="Stratagene mouse embryonic carcinoma (#937317)"
/tissue_type="carcinoma"
/dev_stage="embryonic"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. P19 cell
line. Average insert size: 1.0 kb; Ori-2AP Xa Vector; -5'
adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor
sequence: 5' CCGAGATTTTATTTTATTTT 3'."
BASE COUNT 159 a 145 c 105 g 88 t
ORIGIN

alignment_scores
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block
US-09-327-750D-12 x AA215070
Align seg 1/1 to: AA215070 from: 1 to: 557
1 MetAlaAnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17

```

484 TGATGAGTCTGCCTTATGCCT 505

seq_name: gb_est2:BC088461

seq_definition: 599 bp mRNA EST 26-JAN-2001
BC088461 NIA mouse 15K cDNA Clone set Mus musculus cDNA clone
H3153D07.5, mRNA sequence.

accession: BG088461
version: 1 GI:12571025
keywords: EST.
source: house musculus.
organism: Mus musculus.
taxonomy: Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 599)
T.S., Carter, M.G., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
Kargul, G.J. Xue, K. et al. 2001
Verification and initial annotation of NIA mouse 15K cDNA clone set
Other-ESTs: H3153D07-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: ckd@nigun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://grc.nia.nih.gov/cDNA/15K.html> for details.
Plasmids: H3153
Seq primer: -21M13 Reverse
High quality sequence stop: 599

```

BASE COUNT      163 a    176 g    176 g    176 g
ORIGIN

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref=taxon:10090
/clone="H3153007"
/clone_lib="NIA Mouse 15K CDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"

note=Vector: pSPOR1; Site_1: SalI; Site_2: NotI; This clone is being rearrayed set of 15,747 clones from 11 different mice. The library was constructed by fusing embryos from unfertilized egg to blastocyst embryos part of E7.5 embryos, extraembryonic part of E7.5 embryos and E12.5 female Mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT )-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression during embryonic development of the mouse embryo, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

alignment_scores:   688.00
                   QValue: 5.548
                   Ratio: 5.548
Percent Similarity: 100.000 Percent Identity: 98.387

Alignment Block:
```

51 GlyGlnIleArgLeuLeuAlaProAsnPheArgTTPalIleProAsnAr 67
| | | | | | | | | | | | | | | | | | | | | | | | | |
157 GGCAGCGCTGCCGACTGCCCTCACTTCGGATGGCCATTCCACACAG 206
| | | | | | | | | | | | | | | | | | | | | | | | | |
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAspMetGluMetPheM 84
| | | | | | | | | | | | | | | | | | | | | | | | | |
207 GCAGATGAATCAGCGGTGGTGAGATGGAGATGATGGAAATGTCCA 256
| | | | | | | | | | | | | | | | | | | | | | | | | |
84 eGIuGlUmetArGgLIuIleArgArgLyLeuArgGluLeuGlnLeuArg 100
| | | | | | | | | | | | | | | | | | | | | | | | | |
257 TGGAGAGATGAGAGAAGATCCGGAAGAAGCTTAGGAGGCTACAGCTGAGA 306
| | | | | | | | | | | | | | | | | | | | | | | | | |
101 AsnCysLeuArgTlleLeuMetGlyLiuLeuSerAanHisAphisPhisH 117
| | | | | | | | | | | | | | | | | | | | | | | | | |
307 AATTGTGTACGATCCTTTATGGGGAGCTGTCTAACACCACCATCACCA 356

117 saspglu^{ph}hec^{ys}leu^{met}pro 124
|||||
357 TGATGAATTCGCTTATGCCT 378

seq_name: gb_est1:AW476468	
seq_documentation_block:	
LOCUS AW476468	628 bp
	mRNA
	EST
	24-FEB-2000

DEFINITION
uq76d02.y1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:2937027 5'
similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74

(HUMAN); mRNA sequence.
 AW476468
 AW476468.1 GI:7046574
 ACCESSION
 VERSION

EST.	house mouse.
KEYWORDS	
SOURCE	Mus musculus
ORGANISM	Mus musculus

REFERENCE 1 (bases 1 to 528)
Eukaryota; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap..>
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

UNPUBLISHED (1997)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cgsabs-remail.nlh.gov

Tissue Procurement: Gilbert Smith, Ph.D.
CUNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLN)
 Funding distribution: NCI-CCAP clone distribution information can be

found through the I.M.A.G.E. Consortium, at:
www.linl.gov/bbr/image/image.htm

Seq. primer: -40RP from Gibco
HL53 quality sequence stop: 456.

FEATURES	source
location/Qualifiers	1. .628
/organism="Mus musculus"	

```
/strain="C25CH 11"
/db_xref="taxon:10090"
/clone="IMAGE:293702"
```

```
/clone_id="NCI_CGAP_G033"  
/tissue_type="pooled lung tumors"  
/lab_host="DH10B (phage-resistant)"
```

/notes-Organ: lung; vector: p1/350-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled

lung tumors with a Not I - *hprt* (ur) primer 13
TGTTACCAATCGAATGGGAGCGGCCGCTCTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors.

(Pharmacia), digested with *NotI* and cloned into the *NotI* and *EcoRI* sites of the modified pT73 vector. Library went through one round of normalization, and was

constructed by SENCIO SORRES and M. FALLMA BOMARDU.

seq_name: gb_est1.AA253897

seq_documentation_block:

LOC AA253897 468 bp mRNA
DEFINITION myeloid_r1 gerstaecker mouse pooled organs MRLRR4 Mus musculus cDNA
NOTE similar to G5:M8188 OVARIAN GRANULOSA CELL
CDS CDS PROTEIN RS074 (HUMAN); mRNA sequence.

ACCESSION AA253897
VERSION AA253897.1 GI:1088253
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
TAXIDON 10090
Eukaryota; Chordata; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eumammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 468)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubucque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Thompson,E.
TITLE The Mouse ENVI Mouse EST Project
UNPUBLISHED (1996)
COMMENT Project Comment: Marra M/Mouse EST Project

101 AaccysleuAr9ileuMetGlyLeuSerAanHisAapH1h1 117
|||||
478 AATGTGTCATCCATCTTATGGGGAGCTGTCTAACCCACGATCACC 527
117 sAgGluPhcCysLeuMetPro 124
528 TGTGGAATCTGCTTATGCTT 549
seq_name: gb_est1:A1152323
seq_documentation_block: 430 bp mRNA EST 30-SEP-1998
LOCUS A1152323
DEFINITION ud79a02.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1477034 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION A1152323
KEYWORDS EST 2323.1 GI:368072
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Norris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theisinger, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
The Washington University School of Medicine
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
MGI:94292
Seq primer: -28m3 rev2 ET from Amersham
High quality sequence stop: 419.
Location/Qualifiers
1. 430
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1477034"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH108"
Note: Vector: p7730-Pac (Pharmacia) with a modified
polylinker. 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified p773 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo.
BASE COUNT 135 a 111 c 113 g 71 t
ORIGIN
source
alignment_scores:
Quality: 681.00 Length: 124
Ratio: 5.517 Gaps: 0
Percent Similarity: 99.194 Percent Identity: 97.581
alignment_block:
US-09-327-750d-12 x A1152323
Align seg 1/1 to: A1152323 from: 1 to: 430

1 MetAaAnValH1eGInGluAaSnGluGluMetGluGInProLeuGlnAa 17
|||||
54 ATGCCCAATGTCCACCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 103
17 nGlyGluGluAaSPaGProValGlyGlyGlyGlyGlyGlyGlyGlyGly 34
|||||
104 TGGACAG 153
34 LyAaAnAaAnAaAnAaAnAaAnAaAnAaAnAaAnAaAnAaAnAaAnAa 50
|||||
154 CAAC 203
51 GlyGlnAa 67
|||||
204 GCCCAGGCTCCCGAGCTTCCCTTAACCTCCGATGGGCCATTCCCAAC 253
67 GcInMetAaAnSPGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGly 84
|||||
254 GCAGATGANTGACGGGTGGGTGGAGATGAGATGATGGAATGCTCA 303
84 eGluGluMetAaAaGluAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 100
|||||
304 TGGAGGAGATGAGAGAGATCCCGAAGAAAGCTTAGGGAGCTACAGCT 353
101 AaccysleuAr9ileuMetGlyLeuSerAanHisAapH1h1 117
|||||
354 AATGTGTCATCCATCTTATGGGGAGCTGTCTAACCCACGATCACC 403
117 sAgGluPhcCysLeuMetPro 124
404 TGTGGAATCTGCTTATGCTT 425
seq_name: gb_est1:AW908751
seq_documentation_block: 503 bp mRNA EST 25-MAY-2000
LOCUS AW908751
DEFINITION uf57a05.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1515440 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION AW908751
KEYWORDS EST 80751.1 GI:8073984
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Other ESTs: uf57a05.xl
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
MGI:94292
Seq primer: -40rp from Gibco
High quality sequence stop: 468.
Location/Qualifiers
1. 503
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1515440"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/lab_host="mammary gland"
/lab_host="DH108"
Note: Vector: p7730-Pac (Pharmacia) with a modified
polylinker. 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I

[illegible]


```
repeat_region /clone_lib-"C19785K-A2"
109..136
/note-"14 copies 2 mer ac 100 conserved"
repeat_region /note-"14 copies 2 mer ac 100 conserved"
183..258
/note-"LIMS repeat: matches 7840..7913 of consensus"
misc_feature /177..146
/note-"match: GSS: Em:AQ373724"
repeat_region /1301..1356
/note-"L2 repeat: matches 2693..2750 of consensus"
repeat_region /2568..5038
/note-"L1MD1 repeat: matches 3673..6170 of consensus"
repeat_region /5036..5486
/note-"L1Mc repeat: matches 2052..2207 of consensus"
repeat_region /5496..6885
/note-"L1MD repeat: matches 1..1538 of consensus"
repeat_region /6896..7185
/note-"L2 repeat: matches 2356..2710 of consensus"
misc_feature /complement(7185..7611)
repeat_region /7369..7591
/note-"match: GSS: Em:AQ566001"
misc_feature /complement(7113..8092)
repeat_region /7904..7978
/note-"MIR repeat: matches 34..256 of consensus"
repeat_region /8086..8300
/note-"MIR repeat: matches 20..95 of consensus"
repeat_region /8354..8565
/note-"MIR repeat: matches 8..255 of consensus"
misc_feature /8549..8864
/note-"match: GSS: Em:AQ059599"
repeat_region /8871..9287
/note-"L1MD repeat: matches 1..280 of consensus"
misc_feature /8897..9290
/note-"match: GSS: Em:AQ343214"
repeat_region /9235..9361
/note-"L2 repeat: matches 2619..2749 of consensus"
misc_feature /9378..9890
/note-"match: GSS: Em:AQ284812"
repeat_region /10135..10428
/note-"AluX repeat: matches 1..294 of consensus"
repeat_region /10429..10563
/note-"AluX repeat: matches 1..135 of consensus"
repeat_region /10811..10888
/note-"L1MD repeat: matches 2648..2705 of consensus"
repeat_region /10818..11066
/note-"MIR repeat: matches 2..255 of consensus"
repeat_region /1112..11324
/note-"MIR repeat: matches 11..255 of consensus"
repeat_region /11869..12012
/note-"MIR repeat: matches 20..167 of consensus"
repeat_region /12925..13223
/note-"AluY repeat: matches 1..300 of consensus"
misc_feature /13336..13383
/note-"MIR repeat: matches 9..214 of consensus"
repeat_region /13688..13719
/note-"MIR repeat: matches 77..139 of consensus"
repeat_region /13840..13966
/note-"MIR repeat: matches 22..147 of consensus"
repeat_region /14162..14245
/note-"MIR repeat: matches 46..136 of consensus"
repeat_region /14925..15058
/note-"MIR repeat: matches 48..191 of consensus"
repeat_region /15994..16040
/note-"MIR repeat: matches 102..144 of consensus"
repeat_region /16018..16188
/note-"MIR repeat: matches 3..70 of consensus"
repeat_region /16897..17005
/note-"MIR repeat: matches 9..118 of consensus"
repeat_region /17618..17700
/note-"MIR repeat: matches 108..192 of consensus"
repeat_region /17748..17912
/note-"MIR repeat: matches 73..245 of consensus"
```

BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN HSLRNVSTDPHHHDEFCCLMP*

alignment_scores:
Quality: 194.50 Length: 113
Ratio: 2.739 Gaps: 3
Percent Similarity: 62.832 Percent Identity: 43.363

alignment_block:
US-09-327-750d-13 x AF220189 ..

Align seg 1/1 to: AF220189 from: 1 to: 828

```
1 MetAlaAsnIleHisGlnGluAngluMetGluInProMetGlnAs 17
||| ||| :|||:|||||:||||| :|||: :|||: :|||:
206 ATGGAAATGCCAACCAAGAAATGAGAAAGAGCAAGTCTTAATAA 255
17 ndlyGluGlu...AspArgProLeuGlyGlyGluGlyHisGlnProA 33
:||||| :||||| :||||| :||||| :||||| :|||||
256 AGGGAGGAGCCCTGGCCCTCTTGGATGCTGTGTATCTGCTGCTTA 305
33 IsGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPheArg 48
||||| :||||| :||||| :||||| :||||| :|||||
306 GAGGAATCTAGCGGCTTCGCGTTAGCGCCATCTCTGCAGTATAGA 355
49 TrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAs 65
||| :||| :||| :||| :||| :|||
356 TGGATATGATGATAGCTTGGAGTGGAGAACACAGCAAGGATGAGAAGA 405
65 pApMetGluIlePheMetGluGluMetArgIuIleArgArgLysLeuA 82
:||||| :||||| :||||| :||||| :||||| :|||||
406 GATATGCGAAGCATGGGAGGAGGTGACACGCTGTGGAAAGCTGA 455
82 TgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSer 98
||||| :||||| :||||| :||||| :||||| :|||||
456 GGGAAAGACGTTGAGTATAGTCTCGGGCAGTCACACTGAC...CCC 502
99 AanHisHisAspHisHisAspLupPheCysLeuMetPro 111
||||| :||||| :||||| :||||| :||||| :|||||
503 CCTCACCATGACCATCATGATGAGTTTGGCTTATGCC 541
```

[illegible]

Align seg 1/1 to: N34237 from: 1 to: 519

human

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/resources.shtml
Seq primer: -400P from Gibco
High quality sequence stop: 422.

FEATURES

source
1..612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CGAP-lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site:1:
Salt; Site:2: NotI; Cloned unidirectionally. Primer:
oligo dt. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"
BASE COUNT 169 a 145 c 120 g 178 t
ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-13 x AW512400/rev ..
Align seg 1/1 to reverse of: AW512400 from: 1 to: 612
1 MetAlaAGnIleHIScInGluAsnGluGluMetGluGlnProMetGlnAs 17
581 ATGCCAAATATTCACGAGAAAGAGAGATGGAGCAGCTATGCAGAA 532
17 nGlyGluGluAsnArgProLeuGlyGlyGlyGlyHisGlnProHlaG 34
531 TGGAGAGGAGACCCCTTTGGGAGGAGGTGAAGGCCACCGCTGCAG 482
34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
481 GAAATCGAGGGGACAGCCTCCCGACTTGCCTTAATTTTCGATGGCC 432
51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
431 ATACCAATAGCCATCATCATGATGGATGGGTGGAGATGAGATGATAT 382
67 tGluIlePheMetGluGluMetArgGluIleArgArgLeuArgGlu 84
381 GGAATATTTCATGGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 332
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
331 TGCAGTTGAGGAATTTGTCGCTATCTTATGGGGAGCTCTCTAATCAC 282
101 HisAspHisHisAspGluPheCysLeuMetPro 111
281 CATGACCATCATGATGAATTTGGCTTATGCTT 249

seq_name: gb_est1.A1193112
seq_documentation_block:
LOCUS A1193112 615 bp mRNA EST
DEFINITION Q65909.x1 Soares_fetal_lung_NDHL19W Homo sapiens
IMAGE:1744288.3' similar to gb:M3188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HCR74 (HUMAN); mRNA sequence.
ACCESSION A1193112
VERSION A1193112.1 GI:3744321
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 615)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Soares_fetal_lung_NDHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: p773D (Pharmacia) with a
modified EcoRI linker; Site:1: NotI; Site:2: EcoRI;
15-TGTTACCAATCAGTGGAGCGGCGCAATTTTATT-3';
double-stranded cDNA was size selected, ligated to EcoRI
adapters (Pharmacia), digested with NotI and cloned into
the NotI and EcoRI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Ronaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NDHL19W."
BASE COUNT 169 a 145 c 121 g 180 t
ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-13 x A1193112/rev ..
Align seg 1/1 to reverse of: A1193112 from: 1 to: 615
1 MetAlaAGnIleHIScInGluAsnGluGluMetGluGlnProMetGlnAs 17
584 ATGCCAAATATTCACGAGAAAGAGAGATGGAGCAGCTATGCAGAA 535
17 nGlyGluGluAsnArgProLeuGlyGlyGlyGlyHisGlnProHlaG 34
534 TGGAGAGGAGACCCCTTTGGGAGGAGGTGAAGGCCACCGCTGCAG 485
34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
484 GAAATCGAGGGGACAGCCTCCCGACTTGCCTTAATTTTCGATGGCC 435
51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
434 ATACCAATAGCCATCATCATGATGGATGGGTGGAGATGAGATGATAT 385
67 tGluIlePheMetGluGluMetArgGluIleArgArgLeuArgGlu 84
384 GGAATATTTCATGGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 335
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
334 TGCAGTTGAGGAATTTGTCGCTATCTTATGGGGAGCTCTCTAATCAC 285
101 HisAspHisHisAspGluPheCysLeuMetPro 111

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 image.lnl.gov/image/hcm1/iresources.shtml
 Seq primer: 400p from Gibco
 High quality sequence stop: 422.

FEATURES

source
 1. 612
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2849369"
 /clone_lib="NCI-CGAP_Lym12"
 /tissue_type="lymphoma, follicular mixed small and large
 cell"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: pCMV-SPORT6; Site: 1:
 Note: Site 2: Not; Cloned unidirectionally. Primer:
 Oligo: 81: 11547-015
 catalog #: 11547-015
 BASE COUNT 169 a 145 c 120 g 178 t
 ORIGIN

alignment_scores
 Quality: 609.00 Length: 111
 Ratio: 5.486 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AWS12400/rev ..

Align seg 1/1 to reverse of: AWS12400 from: 1 to: 612

1 MetAlaAsnIleHisGlnGluAngluMetGluGlnProMetGlnAs 17
 581 ATGGCAATATTCACCAAGAAACAGAGAGTGGAGCGCTATGCAGAA 532
 17 nGlyGluGluAspArgProLeuGlyGlyGlyHisGlnProAlaG 34
 531 TGGAGAGAGACCGCCCTTTGGGAGAGGTGAAGCCACCGCTGCAG 482
 34 lYAspArgGlyGlyGlnIleArgGlyLeuAlaProAsnPheArgTrpAla 50
 481 GAAATCGACGGGACAGCGCTCCGCACTTCCCTTATTTTCGATGGGCC 432
 51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
 431 ATACCAATAGGCAGATCAATGATGGATGGTGGGATGGAGATGATAT 382
 67 tGluIlePheMetGluMetGluMetArgGluIleArgArgLysLeuArgGlu 84
 381 GGAATATTCATGGAGAGATGAGAGAAATCAGAGAAACCTTAGGAGC 332
 84 euGluLeuArgAsnCysLeuArgGlyLeuMetGlyGlyLeuSerAsnHis 100
 331 TGCAGTTCAGGAATTCCTGCGTATCTCTTATGGGGAGCTCTCTAATCAC 282
 101 HisAspHisHisAspGluPheCysLeuMetPro 111
 281 CATGACCATCATGATGAATTTTCCTTATGCTT 249

seq_name: gb_est1:AI193112

seq_documentation_block:
 LOCUS AI193112 615 bp mRNA EST 29-OCT-1998
 DEFINITION Q66309.X1 Soares fetal_lung_NHL19W Homo sapiens CDNA clone
 IMAGE:1744288 3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AI193112
 VERSION AI193112.1 GI:374321
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 615)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
 Email: cygbes-f@mail.nih.gov
 Tel: 301-401-9411
 Fax: 301-401-9411
 IMAGE Consortium (info@imga.lnl.gov) for further information.
 Insert Length: 845 Std Error: 0.00
 seq primer: 400p from Gibco
 High quality sequence stop: 445.

FEATURES

source
 1. 615
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1744288"
 /clone_lib="Sources_fetal_lung_NHL19W"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: p773D (Pharmacia) with a
 modified polylinker; Site: 1: Not 1; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGCCCAATTTTTTTTTTT-3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Patricia Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal Heart NHL19W.
 BASE COUNT 169 a 145 c 121 g 180 t
 ORIGIN

alignment_scores
 Quality: 609.00 Length: 111
 Ratio: 5.486 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AI193112/rev ..

Align seg 1/1 to reverse of: AI193112 from: 1 to: 615

1 MetAlaAsnIleHisGlnGluAngluMetGluGlnProMetGlnAs 17
 584 ATGGCAATATTCACCAAGAAACAGAGAGTGGAGCGCTATGCAGAA 535
 17 nGlyGluGluAspArgProLeuGlyGlyGlyHisGlnProAlaG 34
 534 TGGAGAGAGACCGCCCTTTGGGAGAGGTGAAGCCACCGCTGCAG 485
 34 lYAspArgGlyGlyGlnIleArgGlyLeuAlaProAsnPheArgTrpAla 50
 484 GAAATCGACGGGACAGCGCTCCGCACTTCCCTTATTTTCGATGGGCC 435
 51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
 434 ATACCAATAGGCAGATCAATGATGGATGGTGGGATGGAGATGATAT 385
 67 tGluIlePheMetGluMetGluMetArgGluIleArgArgLysLeuArgGlu 84
 384 GGAATATTCATGGAGAGATGAGAGAAATCAGAGAAACCTTAGGAGC 335
 84 euGluLeuArgAsnCysLeuArgGlyLeuMetGlyGlyLeuSerAsnHis 100
 334 TGCAGTTCAGGAATTCCTGCGTATCTCTTATGGGGAGCTCTCTAATCAC 285
 101 HisAspHisHisAspGluPheCysLeuMetPro 111

```

34  lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
502  GAAATCGAGGGAGACGCTGCACATTCGCCCTTAATTTTCGATGGCC 453
51  IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
452  ATACCAATGCGCAGCATATGCGCTGGGTGGATGATGATGATAT 403
67  tGluIlePheMetGluGluMetArgGluIleArgGlyAspGlyAspGlu 84
402  GGAATATTTCATGGAGGAGATGAGAGAAATCAGAGAAATTCAGGGAGC 353
84  euGlnLeuArgAnCysLeuArgIleLeuMetGlyGlyLeuSerAsnHis 100
352  TCGATGTAGGAATGCTTCGCTATCTTATGATGGGAGCTCTCTAATC 303
101  HisAspHisHisAspGluPheCysLeuMetPro 111
302  CATGACCATCATGATGAATTTGCTTATGCT 270
seq_name: gb_est1.A1929106

seq_documentation_block: 662 bp mRNA EST 23-AUG-1999
LOCUS A1929106
DEFINITION au5BD10.Y1 Schneider fecal brain 00004 Homo sapiens cDNA clone
IMAGE2519611.3 similar to J838188 OVARIAN GRANULOSA CELL 13.0
IMAGE2519611.3 similar to J838188 OVARIAN GRANULOSA CELL 13.0
ACCESSION A1929106
VERSION A1929106.1 GI:5665070
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Miller,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S.,
Kliman,D., Kucabinski,M., Lacy,K., deN.T., Lemm,G., Marra,M., Martin,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WASHU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT
Unpublished: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wuston.wustl.edu
This clone is available royalty-free through LNL; contact the
LNL manager (lml@lml.lnl.gov) for further information.
Seq primer: 40BP from Gluc.
High quality sequence stop: 437.
FEATURES
Location/Qualifiers
1..662
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2519611"
/clone_lib="Schneider fetal brain 00004"
/contig="1"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGGAGGAGAGAGCTCAGAGATCTTAATAATTATCCCCCCCCCCCC-3'
5'-GAGGAGGAGCTCAGGATTTTTTTTTTTTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the Cap-trapper method for full-length
transcripts."

```



```

34 lyAenAgaRgLyGlnInlaRgaRgIueUlaProaNgPheArTgTpaIa 50
285 GAATTCGACGGGACGGCTCCGACATTTTCCCTATTTTTCATGGAGCC 334
51 tleProaNaTgGlnIleEnaAspGlyMetGlyGlyAspGlyVaspaPme 67
335 ATACCCAAATGAGCAGACATCAATGATGGGATGGGTGGAGATGGAGATGAT 384
67 tGluIlePheMetGluGluMetArgGdluIleArGaRgLyGlyLeuArgGluL 84
385 GGAATATTCATGGAGGATGAGATGAGAAATCAGAAACAAATTCAGGAGCC 434
84 euGlnLeuRbaRbaCysLeuRgTgTlGleMetGlyGlyLeuSerAenHis 100
435 TGCAGTTGAGGAATGTCGTGCGTATCCTTATGGGGAGCTCTCTATATCAC 484
101 HisAspHisHisAspGluPheCysLeuMetPro 111
485 CATGACCATCATGATGAATTTGGCTATTGCTT 517

seq_name: gb_est2: BG705685

seq_documentation_block: 567 bp mRNA EST 07-MAY-2001
LOCUS BG705685
DEFINITION 502656891.0P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4791506 5',
mRNA sequence.
BUT705685
BUT705685.1 GI:13980274
FEATURES
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 NIH-MGC http://mgc.nci.nih.gov/;
PubMed 119999999
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
UNPUBLISHED 1999
CONTACT: Robert Strausberg, Ph.D.
Email: cagpas@rmail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
Patched: LUMI0688 row: d column: 03
High quality sequence map: 507.
1 Localizer/Qualifiers
1 567
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4791506"
/clone_lib="NIH-MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH108"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site1: BamHI; Site2: SalI-XhoI (gtcggc
AT); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTT-3',
size selected for average insert size 2.3 kb and
average length 807 bp. Library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
160 a 130 c 168 g 109 t

```

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/getchm12.pl?l1=PM04t2-PM0-HT0911-080201-012-h10t3-2001-02-08&t=1)
 Seq primer: puc 18 forward
 High quality sequence (COB: 550).

FEATURES

1..550
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0911"
 /dev_stage="Adult"
 /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORFEST PCR (O.S. Letters Patent application
 no.96/16 to Ludwig Institute for Cancer Research)
 Note: This library was constructed by restriction of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 134 a 113 c 137 g 146 t
 ORIGIN

alignment_scores:
 Quality: 596.00 Length: 112
 Ratio: 5.369 Gaps: 1
 Percent Similarity: 99.107 Percent identity: 99.107

alignment_block:

US-09-327-750D-13 x BG993540

Align seg 1/1 to: BG993540 from: 1 to: 550

```

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
|||||
48 ATGGCAATATTATCCAGGAAACGAAAGAGATGGAGCAGCTATGCAGAA 97
|||||
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
98 TGGAGAGGAGAGCCGCCCTTTGGGAGGAGGTGAAGGCCACCGCCCTCGAG 147
|||||
34 LysAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
|||||
148 GAATCGACGGGACAGGCTCGCCACTTCGCCCTTAATTTTCGATGGCC 197
|||||
50 AtleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspM 67
|||||
198 CATACCAATAGCCAGCATCAATGATGGATGGGTGGAGATGGAGATGATA 247
|||||
67 etGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 83
|||||
248 TGGAAATATTATCGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAG 297
|||||
84 LeuGlnLeuArgAsnCysLeuArgGluLeuMetGlyGlyLeuSerAsnHis 100
|||||
298 CTGCAGTTGAGGAATTTCTTCGCTATCTTATGGGGAGCTCTCTTATCA 347
|||||
100 sHisAspHisHisAspGluPheCysLeuMetPro 111
|||||
348 CCATGACCATCATGATGAATTTTGCCTTATGCTCT 381

```

seq_name: gb_sst1:BE252272

seq_documentation_block: 559 bp mRNA EST 13-JUL-2000
 LOCUS BE252272
 DEFINITION B0411413F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3354819 5',
 accession BE252272
 VERSION BE252272
 KEYWORDS EST.

SOURCE

ORGANISM

human.
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 559)

AUTHORS

NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Genomics 1998;32:1-12

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cga@pds-1.nhlbi.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L2M157 row: n column: 04
 High quality sequence (COB: 559).

FEATURES

source

1..559
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_16"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 15
 adaptor constructs: Library constructed by Ling Hong
 at the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using 2AP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH-MGC Library."

BASE COUNT 157 a 126 c 166 g 110 t
 ORIGIN

alignment_scores:
 Quality: 591.00 Length: 111
 Ratio: 5.422 Gaps: 0
 Percent Similarity: 98.198 Percent identity: 98.198

alignment_block:

US-09-327-750D-13 x BE252272

Align seg 1/1 to: BE252272 from: 1 to: 559

```

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
|||||
177 ATGGCAATATTATCCAGGAAACGAAAGAGATGGAGCAGCTATGCAGAA 226
|||||
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
227 TGGAGAGGAGAGCCGCCCTTTGGGAGGAGGTGAAGGCCACCGCCCTCGAG 276
|||||
34 LysAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
|||||
277 GAATCGACGGGACAGGCTCGCCACTTCGCCCTTAATTTTCGATGGCC 326
|||||
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspM 67
|||||
327 ATACCAATAGCCAGCATCAATGATGGATGGGTGGAGATGGAGATGATAT 376
|||||
67 etGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 84
|||||
377 GGAAATATTATCGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 426
|||||
84 LeuGlnLeuArgAsnCysLeuArgGluLeuMetGlyGlyLeuSerAsnHis 100
|||||
427 TGCAGTTGAGGAATTTCTTCGCTATCTTATGGGGAGCTCTCTTATCA 476
|||||
101 HisAspHisHisAspGluPheCysLeuMetPro 111

```

seq_name	seq	seq_name	seq	seq_name	seq
/cgn2.5/ptodate/2/ina/65	_COMB.seq:US-08-464-514-1 +	70.50	117.16	70.95	230.
/cgn2.5/ptodate/2/ina/68	_COMB.seq:US-08-466-403-1 +	70.50	117.16	70.95	230.
/cgn2.5/ptodate/2/ina/69	_COMB.seq:US-08-485-0180-1 +	70.50	116.54	76.85	79.
/cgn2.5/ptodate/2/ina/5A	_COMB.seq:US-08-332-4678-2 +	70.00	126.70	20.86	797
/cgn2.5/ptodate/2/ina/5A	_COMB.seq:US-08-681-811-2 +	70.00	126.70	20.86	797
/cgn2.5/ptodate/2/ina/5B	_COMB.seq:US-08-387-942C-18				

seq documentation block:
: Sequence 18, Application US/08387942C
: Patent No. 593289
: GENERAL INFORMATION:
: APPLICANT: ERTESVAG, HELGA
: APPLICANT: VALLA, SVEN
: APPLICANT: SKJAK, BRENN
: APPLICANT: LARSEN, BJORN
: TITLE OF INVENTION: COMPOUNDS COMPRISING SOURCES
: INCLUDING MANNOSE
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:

ADDRESSSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
 STREET: P.O. BOX 747
 CITY: FALLS CHURCH
 STATE: VA
 COUNTRY: USA
 ZIP: 22042
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 SOFTWARE: PC DOS/MS-DOS
 CURRENT APPLICATION DATA: Version #1.30
 CURRENT APPLICATION DATA: Release #1.0, Version #1.30

```

SOFTWARE: PatentIn Release 11.0, Ver
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387.942C

```

COMMUNICATION NUMBER: 08/387,942C
 FILING DATE: 09-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURPHY JR, GERALD M.
 REGISTRATION NUMBER: 28,977
 REFERENCE/DOCKET NUMBER: 1809-106P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-205-8000
 TELEFAX: 703-205-8050
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1155 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 05-08-387-942C-18

alignment_scores:		
Quality:	88.50	Length: 79
Ratio:	2.011	Gaps: 4
Percent Similarity:	55.696	Percent Identity: 36.709

```
alignment_block:
US-09-327-750D-13 x US-08-387-942C-18
```

Align seg 1/1 to: US-08-387-942C-18 from: 1 to: 1155

18 GlyGluGlnAspArgProLeuGlyGlyGlyGlyHisGlnProAlaGI 34
|||||::: ||||| ||||| :: :: |||||

72 GGCGATGACAGCCGGCCCCCACCGCGCGCGCGCGCTACCTGC CGCG
34 YAGNArg... ArgGLVGLNALAARGARGLLeuA!aProAsnpheArgTTPA 50

122 CCGCGAATATCGGTTCAGCGCGCGAGGAGCC....TTCGATGGT 165

50 laile.....ProAsnArgGlnIleAsnAspGlyMet 60
||| |||:: ::|||

100 TGTCTGACCATCAAGAGAGCAACGTCCTCCATATCGTCGGCGCCGGGATGGGCGA 215

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: alternative 3' end
INDIVIDUAL ISOLATE: starting at nt. 3012 of SEQ ID NO:9
FEATURE:
NAME/KEY: CDS
LOCATION: 2..1021
PCT-US96-02331-12

alignment_scores:
Quality: 82.00 Length: 125
Ratio: 1.302 Gaps: 4
Percent Similarity: 50.400 Percent Identity: 20.800

alignment_block:
US-09-327-750D-13 x PCT-US96-02331-12

Align seg 1/1 to: PCT-US96-02331-12 from: 1 to: 1244

5 HisGlnGluAsnGluGluMetGluGlnProMetGlnAsnGlyGluGluAs 21
|||||
26 CAGCAGCGTATCCGGATCGGAAGTCTGATGCAATCATCGGAGCAGCA 75
|||||
21 p.....ArgProL 24
76 CGACGATCCAGCGCTTATCGAGGAGTCTGTGATGATCACGTTCTGTGAGA 125
24 euGlyGlyGluGluGluHisGlnProAlaGlyAsnArgArgGlyGlnAla 40
|||||
126 TCGAGCGGGAGTACGACGATCCGGAGGAGATGAAGGAGGAGCAGCTAC 175
41 ArgArgLeuAlaProAsnPhaArgTrpAlaIle.....ArgProL 24
176 CATGCCACACCGCCCAAGTACAGACGGCTGTGTTTATGCTCTCCGCA 225
52 .ProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetG 68
226 TCGGATGAAGAGCGGCTCCGGATCGGATCGGAT..... 262
68 LuIlePheMetGluGluMetArgGluLeuArgGlyLeuArgGluLeu 84
|||||
263 ..ATCTATGATGCGCGCTGCTGATGCAATTCGAGTACAGTGAAGAGCTC 310
85 GlnLeuArgAsnCysLeuArgIle.....LeuMetGlyGly 96
311 AACATGCGCGCAACATGATGATGATGATGATGATGATGATGATGATG 360
96 uLeuSerAsnHisHisAspHisHis 104
361 CTATTCCCGCATCATCGGACCAT 385

seq_name: /usr2_6/ptodate/2/ins/PTCUS_COMB.seq:PCT-US96-02331-14

seq_documentation_block:
GENERAL INFORMATION:
APPLICANT: The Board of Trustees of the Leland Stanford Junior
APPLICANT: University
TITLE OF INVENTION: Board of Regents, The University of Texas System
TITLE OF INVENTION: Methods and Compositions for Altering
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/02331

FILING DATE: 09-FEB-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/386,495

FILING DATE: 10-FEB-1995

ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0950

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4255 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: cDNA to mRNA

REPERES: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: fruitless transcript in Fig. 7E

FEATURE:

NAME/KEY: CDS

LOCATION: 1507..4032

PCT-US96-02331-14

alignment_scores:
Quality: 82.00 Length: 125
Ratio: 1.302 Gaps: 4
Percent Similarity: 50.400 Percent Identity: 20.800

alignment_block:
US-09-327-750D-13 x PCT-US96-02331-14

Align seg 1/1 to: PCT-US96-02331-14 from: 1 to: 4255

5 HisGlnGluAsnGluGluMetGluGlnProMetGlnAsnGlyGluGluAs 21
|||||
3037 CAGCAGCGTATCCGGATCGGAAGTCTGATGCAATCATCGGAGCAGCA 3086
|||||
21 p.....ArgProL 24
3087 CGACGATCCAGCGCTTATCGAGGAGTCTGTGATGATCACGTTCTGTGAGA 3136
24 euGlyGlyGluGluGluHisGlnProAlaGlyAsnArgArgGlyGlnAla 40
|||||
3137 TCGAGCGGGAGTACGACGATCCGGAGGAGATGAAGGAGGAGCAGCTAC 3186
41 ArgArgLeuAlaProAsnPhaArgTrpAlaIle..... 51
3187 CATGCCACACCGCCCAAGTACAGACGGCTGTGTTTATGCTCTCCGCA 3236
52 .ProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetG 68
3237 TCGGATGAAGAGCGGCTCCGGATCGGATCGGAT..... 3273
68 LuIlePheMetGluGluMetArgGluLeuArgGlyLeuArgGluLeu 84
|||||
3274 ..ATCTATGATGCGCGCTGCTGATGCAATTCGAGTACAGTGAAGGAGCTC 3321
85 GlnLeuArgAsnCysLeuArgIle.....LeuMetGlyGly 96
3322 AACATGCGCGCAACATGATGATGATGATGATGATGATGATGATGATG 3371
96 uLeuSerAsnHisHisAspHisHis 104
3372 CTATTCCCGCATCATCGGACCAT 3396

OM of: US-09-327-750D-13 to: N_Geneseq_1101.* out_format : pfs

Date: Mar 11, 2002 3:42 PM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame; 22m model -DRV=xlh
-O=Cvnt2.1\USPTO.spool\US09327750\runat_11032002_101154_20362/app_query.fasta.1.1472
-DB=N_Geneseq_1101 -OFM=fastcap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-TRANS=human40.cdi -LIST=45 -DOCLIM=200 -THR_SCORE=pcr
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs
-USPR=US09327750 -ENV=1 -MINLEN=0 -MAXLEN=200000000
-DEVT=US09327750 -ENV=1 -MINLEN=0 -MAXLEN=200000000
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_AUPFI -WAIT_THREADS=1

Search information block:

Query: US-09-327-750D-13
Query length: 111
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:

Sequence	Strd Orig	2Score	EScore	Len	Documentation
/SID2/gcgdata/geneseq/NA2001.DAT.AAF23529 +		609.00	1137.76	4	4e-58
/SID2/gcgdata/geneseq/NA2001.DAT.AAF23529 +		582.00	1134.20	4	4e-55
/SID2/gcgdata/geneseq/NA2001.DAT.AAF23529 +		564.50	1102.28	2	2e-53
/SID2/gcgdata/geneseq/NA2001.DAT.AAC01005 +		537.00	1050.66	1	1e-50
/SID2/gcgdata/geneseq/NA2001.DAT.AAC01004 +		521.00	1019.98	9	9e-49
/SID2/gcgdata/geneseq/NA2001.DAT.AAH03517 +		277.00	532.88	1	1e-21
/SID2/gcgdata/geneseq/NA2001.DAT.AAH3750 +		277.00	529.49	2	2e-21
/SID2/gcgdata/geneseq/NA2001.DAT.AAC95548 +		277.00	528.48	2	2e-21
/SID2/gcgdata/geneseq/NA2001.DAT.AAH87147 +		193.00	381.92	3	3e-13
/SID2/gcgdata/geneseq/NA2001.DAT.AAH5343 +		187.50	357.12	8	8e-12
/SID2/gcgdata/geneseq/NA2001.DAT.AAH75810 +		175.00	332.76	1	1e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAH28581 +		175.00	331.63	2	2e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAH38581 +		165.00	321.23	2	2e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAC03880 +		165.00	321.23	7	7e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAC03880 +		169.00	321.91	7	7e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAC03879 +		169.00	321.81	7	7e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAH8286 +		168.00	324.63	5	5e-10
/SID2/gcgdata/geneseq/NA2001.DAT.AAH60367 +		162.00	306.02	5	5e-09
/SID2/gcgdata/geneseq/NA2001.DAT.AAC01003 +		108.00	206.81	0	0.0159
/SID2/gcgdata/geneseq/NA2001.DAT.AAH72409 +		105.00	190.32	0	0.0159
/SID2/gcgdata/geneseq/NA2001.DAT.AAH14082 +		102.50	186.05	0	0.0076
/SID2/gcgdata/geneseq/NA2001.DAT.AAH135463 +		102.50	186.05	0	0.0076
/SID2/gcgdata/geneseq/NA2001.DAT.AAH135463 +		91.50	177.35	0	0.0819
/SID2/gcgdata/geneseq/NA2001.DAT.AAH14628 +		91.50	177.53	0	0.0819
/SID2/gcgdata/geneseq/NA2001.DAT.AAH107131 +		91.50	163.90	0	0.0716
/SID2/gcgdata/geneseq/NA2001.DAT.AAH68557 +		91.50	163.26	0	0.119
/SID2/gcgdata/geneseq/NA2001.DAT.AAC98961 +		91.50	162.81	0	0.5421
/SID2/gcgdata/geneseq/NA2001.DAT.AAH59746 +		88.50	161.17	0	0.5689
/SID2/gcgdata/geneseq/NA2001.DAT.AAH63293 +		87.50	155.01	19	19.17
/SID2/gcgdata/geneseq/NA2001.DAT.AAH81459 +		87.50	117.80	174	174.20
/SID2/gcgdata/geneseq/NA2001.DAT.AAF21610 +		87.50	100.90	1	1e+03
/SID2/gcgdata/geneseq/NA2001.DAT.AAH81490 +		87.50	87.24	8	8e+03
/SID2/gcgdata/geneseq/NA2001.DAT.AAH19956 +		84.50	151.53	2	2e-03
/SID2/gcgdata/geneseq/NA2001.DAT.AAH24631 +		82.50	128.35	45	45.02
/SID2/gcgdata/geneseq/NA2001.DAT.AAH71031 +		82.00	122.67	25	25.88
/SID2/gcgdata/geneseq/NA2001.DAT.AAH71031 +		80.00	146.02	6	6.57
/SID2/gcgdata/geneseq/NA2001.DAT.AAH71031 +		80.00	140.62	4	4.34
/SID2/gcgdata/geneseq/NA2001.DAT.AAC64101 +		80.00	140.62	9	9.34

/SID2/gcgdata/geneseq/NA2001.DAT.AAD04966 + 80.00 140.62 9.34
/SID2/gcgdata/geneseq/NA2001.DAT.AAD04966 + 80.00 140.62 9.34
/SID2/gcgdata/geneseq/NA2001.DAT.AAC64039 + 80.00 140.40 9.61
/SID2/gcgdata/geneseq/NA2001.DAT.AAD04964 + 80.00 140.40 9.61
/SID2/gcgdata/geneseq/NA2001.DAT.AAX41328 + 79.00 149.09 3.15
seq_name: /SID2/gcgdata/geneseq/NA2001.DAT.AAF23529

seq_documentation_block:
10 AAF23529 standard; DNA; 891 BP.
AC AAF23529;

DT 22-MAR-2001 (first entry)
DE Human NADE DNA.

KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.
OS Homo sapiens.

PN WO2000075278-A2.
PD 14-DEC-2000.

PF 07-JUN-2000; 2000WO-US15621.
PR 07-JUN-1999; 99US-0327750.
PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Sato T;
DX WPI; 2001-061707/07.

PT New p75-neurotrophin receptor-associated cell death executor (NADE) and
the gene encoding NADE, useful for modulating the activity of p75NTR
and for detecting neurodegenerative diseases.

PS Disclosure; Fig 1; 134pp; English.

CC The present invention relates to a purified polypeptide capable of
binding neurotrophin receptor (p75-NTR). The invention is useful for
binding and modulating the activity of p75NTR. The peptide mediates
NGF-induced apoptosis, which plays an important role in neurogenetic
diseases. The peptide of the invention and p75NTR are useful for
inhibiting NF-kappaB activation in a cell or a subject, for inducing
NF-kappaB activity to cleave poly (ADP-ribose) polymerase
and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.

Sequence 891 BP; 251 A; 182 C; 224 G; 224 T; 0 other;

alignment_scores:
Quality: 609.00 Length: 111
Percent Similarity: 100.00 Percent Identity: 100.00
Ratio: 5.486
Gaps: 0

alignment_block:
US-09-327-750D-13 x AAF23529

Align seg 1/1 to: AAF23529 from: 1 to: 891

1 MetAlaAsnIleHISGLIuAnGLuMetGLuInProMetGlnAs 17
312 ATGGCAATATTACACGAGAAACGAGAGATGACGACCTATCGAA 361
17 nGLyGLuAspArgProLeuGLyGLyGLyGLyGLyHISGLInProAlaG 34
362 TGGAGAGAAAGCGCCCTTGGAGAGAGTGAAGCCACCGCTCGAG 411
34 lYAsnArgArgGLyGlnAlaArgArgLeuAlaProAsnPhcArgTrpAla 50

PR 29-JAN-1999; 99US-0118078.
 PA (SEAR) SEARLE & CO G D.
 PI Bunch RP, Curtiss SW, Rodi CP, Morris DL;
 XX WPI: 2000-505977/45.
 XX New nucleic acid encoding a carcinogenic biomarker. Induced by
 PT phenobarbital treatment of rat hepatocytes, useful for identifying
 PT carcinogenic compounds.
 XX Claim 1; Page 73; 24Opp: English.
 XX AA87080 to AA87656 represent nucleic acid sequences (N1) encoding a
 CC carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by
 CC creating rat hepatocytes with phenobarbital. The nucleic acids are
 CC useful for identifying carcinogenic compounds. The nucleic acid molecules
 CC are used to identify and/or primers for detecting or inducing
 CC carcinogenesis, respectively.
 XX Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other;

alignment_scores:
 Quality: 193.00 Length: 35
 Ratio: 5.514 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-327-750D-13 x AA87147

Align seg 1/1 to: AA87147 from: 1 to: 187

77 TLeAArgLyLeuArgGluLeuGlnLeuArgAsnCysLeuArgLeLe 93
 |||||
 2 ATCCGGAGAACCTTAGGGAGCTGCGAGTTCAGAAATTCCTGCGTATCT 51
 93 UmetGlyGluLeuSerAnHshHspHshHspHshHspGluPheCysLeu 110
 |||||
 52 TTTGGGGAGCTCTCTATCATCCACGACCATCATCATGATGATTCGCTTA 101
 110 etPro 111
 |||||
 102 TGGCT 106

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:
 ID AAH45143 standard; cDNA; 792 BP.
 XX AAH45143;
 AC
 XX 07-SEP-2001 (first entry)
 XX Human brain expressed X-linked protein, hbex, coding sequence.
 XX Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;
 KW hbex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;
 KW X-chromosome-binding mental retardation; lissencephalous disease; ss.
 XX Homo sapiens.
 XX W0200140286-A1.
 PN 07-JUN-2001.
 XX 27-NOV-2000; 2000WO-CN00502.
 XX 30-NOV-1999; 99CN-0124179.
 XX (BIOR-1) BIOROAD, GENE DEV LTD SHANGHAI.
 XX

PI Mao Y, Xie Y;
 XX WPI: 2001-397944/42.
 DR P-PSDB: AAB99224.
 XX Isolated human brain-expressed X-linked polypeptide used to diagnose
 PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
 PT deafness and X-chromosome-binding mental retardation
 XX Claim 5; Page 22; 30pp: Chinese.
 XX The present sequence is the coding sequence for a human brain-expressed
 CC X-linked protein (hbex). hbex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, X-chromosome-binding mental retardation and
 CC lissencephalous disease. hbex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hbex coding sequence can be used as primers or probes, or in producing
 CC gene chips or microarrays.
 XX Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:
 Quality: 187.50 Length: 113
 Ratio: 2.679 Gaps: 3
 Percent Similarity: 61.947 Percent Identity: 42.478

alignment_block:
 US-09-327-750D-13 x AAH45143

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetaLAsnIleHshGluAsnGluMetGluGlnProMetClnAs 17
 |||||
 212 ATGGAATATCCACCAAGAAATGAAAGGAGCAAGTTCCTAATAA 261
 17 nGlyGluGlu...AspArgProLeuGlyGlyGlyGlyHsGlnProA 33
 |||||
 262 AGGGAGCCCTTGGCCCTCCCTTGGATGCTGCTGAATACTGTGTGCTA 311
 33 LeClyAsnArgGly...GlnAlaArgArgLeuAlaProAnPheArg 48
 |||||
 312 GAGAAATCTATGACCGTTCGCTTAGCCACCCATCTCTCAGATAGA 361
 49 TrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAs 65
 |||||
 362 TGGATATGATCATAGCTTGGAGAACCAAGCAAGGATGAGAGAAGA 411
 65 pAspMetGluIlePheMetGluGluMetArgGluIleArgArgGlyLeuA 82
 |||||
 412 GAATATGGAAAGGATTTGGGAGGGGGTGGAGACAGCTGATGGAAAGCTGA 461
 82 TrpGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSer 98
 |||||
 462 GGGAAAGCAAGTTGCTATAGTCTCGGGCAGTCAGCACTGAC...CCC 508
 99 AnHshHspHshHspHshHspGluPheCysLeuMetPro 111
 |||||
 509 CTCACCATGACCATCATGATGAGTTTGCCTTTATGCC 547

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810

seq_documentation_block:
 ID AAH75810 standard; cDNA; 767 BP.
 XX AAH75810;
 XX 17-OCT-2001 (first entry)
 XX Human X chromosome linked gene expression protein 14 coding sequence.
 DE Human; X chromosome linked gene expression protein 14; cancer;
 KW

Align seg 1/1 to: AAF59611 from: 1 to: 862

```

3 AsnIleHisGlnGluAsn.....GluGluMetGluGlnPrometG1 16
271 MATGTCACCAAGGAAATGATGAAAAGATGAAAGAGCAAGTGTCTAA 320
16 nAnGlyGluGlu...AspArgProLeuGlyGlyGlyGluGlyHisGlnP 32
321 TAAAGGAGGAGCCCTTGGCCCTACCTTTGATGTTAGTAACTACTGTGTC 370
32 ToALeGlyAsnArgGly...GlnAlaArgArgLeuAlaProAsnPhe 47
371 CTAGAGGAACCGTAGGCGGTTCGGCGTTAGCCAGCCATCCGCTGATAT 420
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspG1 64
421 AGATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 470
64 yAspArgMetGluIlePheMetGluGluMetArgGluIleArgArgGlyS1 81
471 GGAGATATGGAAGGATTTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 520
81 euATGGLuLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
521 TGAGGGAAGAGCAGTGTGATGATGATGATGATGATGATGATGATGAT... 567
98 SerAsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
568 CCCCTCCACCATGACCATGACCATGACCATGATGATGATGATGATGAT 609
seq_name: /SID2/seqdata/geneseq/geneseq/NA2001.DAT:AAF59611
seq_documentation_block:
ID AAF59611 standard; cdna; 898 bp.
AC AAF59611;
XX
XX
24-APR-2001 (first entry)
XX
XX Human cell cycle and proliferation protein CCYPR-22 cdna, SEQ ID NO:76.
XX
XX Cell cycle and proliferation protein; CCYPR; human; agonist;
XX antagonist; gene therapy; detection; gene therapy;
XX developmental disorder; tumour; immune disorder;
XX developmental disorder; cell signalling disorders; and cell proliferative
XX disorders including cancer. Specific examples of these disorders
XX include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
XX diabetes mellitus, disorders of the menstrual cycle and infections
XX caused by bacteria.
XX
OS Homo sapiens.
XX
XX WO200107471-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000WO-US19948.
XX
XX 21-JUL-1999; 99US-0145075.
XX
XX 08-SEP-1999; 99US-0153129.
XX
XX 10-NOV-1999; 99US-0164647.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
XX Azimzai Y, Yang J, Lu DM, Baughn MR, Patterson C, Shah P;
XX WPL: 2001-112727/12.
XX
XX P-PSDB; AAB60474.
XX
XX Human cell cycle and proliferation proteins and polynucleotides are
XX used to treat, diagnose and prevent immune, developmental and cell
XX signaling disorders and cell proliferative disorders including cancer -
XX
```

```

PS
XX
XX Claim 5; Page 181-182; 205pp; English.
XX
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
XX cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
XX CCYPR and agonists of CCYPR are used to treat diseases or conditions
XX associated with decreased expression of functional CCYPR, while CCYPR
XX antagonists are used to treat diseases or conditions associated with
XX overexpression of functional CCYPR. Antisense oligonucleotides and
XX antibodies are used to detect CCYPR. CCYPR itself may be used to detect
XX expression of functional CCYPR. Antisense oligonucleotides and
XX radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
XX compounds e.g., antibodies, oligonucleotides and proteins (receptors)
XX that specifically bind to CCYPR, and in drug screening methods to
XX identify compounds that modulate the activity of CCYPR. CCYPR
XX nucleotides can be used to generate transgenic animal models of human
XX disease, and can be used in gene therapy in target cells with genetic
XX abnormalities with respect to the expression of CCYPR for the
XX treatment or prevention of a disorder associated with CCYPR. CCYPR
XX diseases which can be diagnosed, treated and prevented using
XX CCYPR include immune disorders, tumour, immune disorder,
XX developmental and cell signalling disorders, and cell proliferative
XX disorders including cancer. Specific examples of these disorders
XX include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
XX diabetes mellitus, disorders of the menstrual cycle and infections
XX caused by bacteria.
XX
XX Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;
XX
XX
XX alignment_scores:
XX Quality: 175.00 Length: 114
XX Identity: 2.465 Gaps: 4
XX Percent Similarity: 62.381 Percent Identity: 41.228
XX
XX alignment_block:
XX US-09-327-750D-13 x AAF59611
XX
XX Align seg 1/1 to: AAF59611 from: 1 to: 898
XX
XX 3 AsnIleHisGlnGluAsn.....GluGluMetGluGlnPrometG1 16
XX 299 MATGTCACCAAGGAAATGATGAAAAGATGAAAGAGCAAGTGTCTAA 348
XX 16 nAnGlyGluGlu...AspArgProLeuGlyGlyGlyGluGlyHisGlnP 32
XX 349 TAAAGGAGGAGCCCTTGGCCCTACCTTTGATGTTAGTAACTACTGTGTC 398
XX 32 ToALeGlyAsnArgGly...GlnAlaArgArgLeuAlaProAsnPhe 47
XX 399 CTAGAGGAACCGTAGGCGGTTCGGCGTTAGCCAGCCATCCGCTGATAT 448
XX 48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspG1 64
XX 449 AGATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 498
XX 64 yAspArgMetGluIlePheMetGluGluMetArgGluIleArgArgGlyS1 81
XX 499 GGAGATATGGAAGGATTTGGGAGGAGTGGGAGGAGTGGGAGGAGTGGG 548
XX 81 euATGGLuLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
XX 549 TGAGGGAAGAGCAGTGTGATGATGATGATGATGATGATGATGATGAT... 595
XX 98 SerAsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
XX 596 CCCCTCCACCATGACCATGACCATGACCATGATGATGATGATGATGAT 637
XX
XX seq_name: /SID2/seqdata/geneseq/geneseq/NA2000.DAT:AA03880
XX seq_documentation_block:
XX ID AA03880 standard; cdna; 662 bp.
XX AC AA03880;
XX
XX
```

```
US-09-327-750D-13 x AAC10889 ...
Align seg 1/1 to: AAC10889 from: 1 to: 692

 3 AsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAsnGlyG1 19
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 AATGTCAACCAAGGAAATGATGAAAGAGAGAGAGAGAGAGAGAGAGAG 331
19 uGluAspArgProLeuGly.....GlyGlyGluGlyHisGlnP 32
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
332 TAAAGGGAGGCGCTTGCGCCCTTGGATGTYTGTGTACTGTCTGC 381
32 RoIAeGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPhe 47
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
382 CTAGAGGAATCGTAGCGCGTTCGCGTTAGCGAGCCCATCTGCAGTAT 431
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspG1 64
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
432 AGATGGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 481
64 YAspAspMetGluIlePheMetGluGluMetArgGluIleArgArgLysL 81
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
482 AGGATTTGGAAAGGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531
81 euArgGluLeuGlnLeuArgAsnCysLeuArgGlnIleLeuMetGlyGluLeu 97
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
532 TGAGGGAAAGAGAGTTGAGTCAATAGTCTCGCGGCGAGTCAGCAGTGAC... 578
98 SerAsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
579 CCCCCTCACCATGACCATCATGATGATGATGATGATGATGATGATGATG 620
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sv model

Run on: March 11, 2002, 15:42:51 ; Search time 144 Seconds
(Without alignments)
4167.553 Million cell updates/sec

Title: US-09-327-750d-28
Perfect score: 700
Sequence: 1 acagcgtctccgacgacg.....ggggcttgggtccagtg 700

Scoring table: IDENTITY_NUC
Gapop 10.0, Capext 1.0

Searched: 930621 seqs, 42862619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.1
1: /SID52/gcgdata/geneseq/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	693.6	99.1	700	22	AAF23528
2	307.4	43.9	891	22	AAF23529
3	281.6	40.7	891	22	AAF23530
4	281.6	36.2	532	21	AAH01008
5	122.2	27.5	485	21	AAH01004
6	147.2	21.0	187	21	AAH07147
7	127.2	18.2	426	22	AAH2586
8	80.2	11.5	865	22	AAH03517
9	80.2	11.5	1229	22	AAH13750
10	80.2	11.5	1364	22	AAH5548
11	58	8.3	401	21	AAH01003

12	56.4	8.1	792	22	AAH45143
13	55.8	8.0	767	22	AAH75810
14	55.8	8.0	998	22	AAH75811
15	52.6	7.9	692	21	AAH08860
16	52.6	7.9	692	21	AAH08860
17	55.6	7.9	698	21	AAH08860
18	55.6	7.9	698	21	AAH08860
19	54.2	7.7	862	22	AAH50367
20	50.4	7.2	298	21	AAH58581
21	44.6	6.4	5552	21	AAH06182
22	42.2	6.0	128	21	AAH58109
23	41.4	5.9	830	22	AAH12100
24	41.4	5.9	830	22	AAH12100
25	41.4	5.9	830	22	AAH12100
26	40.8	5.8	381	22	AAH10740
27	40.8	5.8	381	22	AAH10740
28	40.8	5.8	381	22	AAH10740
29	40.8	5.8	381	22	AAH10740
30	40.4	5.8	10732	21	AAH10594
31	39.2	5.6	1635	22	AAH12780
32	39.2	5.6	1635	22	AAH12780
33	39.2	5.6	1635	22	AAH12780
34	39.2	5.6	1635	22	AAH12780
35	39.2	5.6	1635	22	AAH12780
36	39.2	5.6	1635	22	AAH12780
37	38.8	5.5	1452	21	AAH26304
38	38.6	5.5	1452	21	AAH26304
39	38.4	5.5	1452	21	AAH26304
40	38.4	5.5	1452	21	AAH26304
41	38.4	5.5	1452	21	AAH26304
42	37.6	5.4	439	22	AAH12718
43	37.6	5.4	439	22	AAH12718
44	37.6	5.4	439	22	AAH12718
45	37.6	5.4	439	22	AAH12718

ALIGNMENTS

RESULT 1
AAF23528 standard; DNA; 700 BP.
II AAF23528 standard; DNA; 700 BP.
AC AAF23528:
DT 22-MAR-2001 (first entry)
DE Mouse NADE DNA.
KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.
OX Mus sp.
PN HQ200073278-A2.
PD 14-DEC-2000.
PF 07-JUN-2000; 2000MO-US15621.
PR 07-JUN-1999; 99US-0327750.
XX (UNCO) UNIV COLUMBIA NEW YORK.
XX Sato T.
XX WPI; 2001-061707/07.
XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX the gene encoding NADE, useful for modulating the activity of p75NTR
XX and for detecting neurodegenerative diseases -
XX Claim 12; Fig 1; 134pp; English.

Claim 1: Page 581-582; 1299pp; English.

[illegible]

AAH02586

.....

DT 26-JUN-2001 (first entry)

THE UNIVERSITY OF CHICAGO

RESULT 8

DT 26-JUN-2001 (first entry)

AC AAF5810;
 XX DT 17-OCT-2001 (first entry)
 XX DE Human X chromosome linked gene expression protein 14 coding sequence.
 XX KW Human; X chromosome linked gene expression protein 14; cancer;
 XX KW HIV infection; cytostatic; anti-HIV; chromosome X; ss.
 XX OS Homo sapiens.
 XX PN CN1296969-A.
 XX PD 30-MAY-2001.
 XX PF 23-NOV-1999; 99CN-0124078.
 XX PR 23-NOV-1999; 99CN-0124078.
 XX PX (SHAN)-SHANGHAI BORONG GENE DEV CO LTD.
 XX PI Mao Y, Xie Y;
 XX DR WPI: 2001-483897/53.
 XX DR P-PSDB: AAG66407.
 XX PT Polypeptide-human X chromosome linked gene expression protein 14 and
 XX PT polynucleotide for coding said polypeptide.
 XX PS Claim 6; Page 24 (Disclosure); 31pp; Chinese.
 XX CC The present sequence is the coding sequence for human X chromosome linked
 XX CC gene expression protein 14. The protein and coding sequence are useful
 XX CC for treating diseases e.g. cancer and HIV infection.
 XX SQ Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;

Query Match 8.0%; Score 55.8; DB 22; Length 767;
 Best Local Similarity 53.7%; Pred. No. 8.4e-06;
 Matches 139; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

OY 316 accaccgaagggccagctgcgcgactgcctcccttaacttcgactggccattcccaaca 375
 DB 255 accctggcgtttccgtttggcagccatcttcgcgtatagttggatataatgagata 314
 OY 376 ggcagatgaatgacgggttggtggagatgagatgataatgaaatgtctatggagaga 435
 DB 315 ggtctggagagccacagggcaaggatgagagagagataatggaaggattggggaggagg 374
 OY 436 tgagagagatccggagaaagattagggagctacagctgagaaattgtcagcatcctta 495
 DB 375 tgagacagctgagaaagagctgggaaagagcgttgagtgatcgtctggcggcagtica 434
 OY 496 tgggggagctgtataaccac 555
 DB 435 gcaatcaccctccctccacatgacacagagagattgcttattgcttattgcttattgct 491
 OY 556 ggtcattccctccctgagat 574
 DB 492 tgatggttcccttgaggt 510

RESULT 14
 AAF59611 standard; cDNA; 898 BP.
 XX AAF59611;
 XX DT 24-APR-2001 (first entry)
 XX DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.

KW Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection; ss.
 XX OS Homo sapiens.
 XX PN NC0200107471-A2.
 XX PD 01-FEB-2001.
 XX PF 21-JUL-2000; 2000MO-US19948.
 XX PR 21-JUL-1999; 99US-0145075.
 XX PR 08-SEP-1999; 99US-0153129.
 XX PX 10-NOV-1999; 99US-0164647.
 XX PA (INCYTE) INCYTE GENOMICS INC.
 XX PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
 XX PI Azimzal Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
 XX DR WPI: 2001-112727/12.
 XX DR P-PSDB: AAB60474.
 XX PT Human cell cycle and proliferation proteins and polynucleotides are
 XX PT used to treat, diagnose and prevent immune, developmental and cell
 XX PT signalling disorders and cell proliferative disorders including cancer.
 XX PS Claim 5; Page 181-182; 205pp; English.

CC Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
 CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in assays for the ability of
 CC CCYPR to bind to CCYPR, and in assays for the ability of CCYPR
 CC nucleotides can be used to generate transgenic animal models of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC Diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders, and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.

SQ Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

Query Match 8.0%; Score 55.8; DB 22; Length 898;
 Best Local Similarity 53.7%; Pred. No. 9e-06;
 Matches 139; Conservative 0; Mismatches 117; Indels 3; Gaps 1;

OY 316 accaccgaagggccagctgcgcgactgcctcccttaacttcgactggccattcccaaca 375
 DB 408 accctggcgtttccgtttggcagccatcttcgcgtatagttggatataatgagata 467
 OY 376 ggcagatgaatgacgggttggtggagatgagatgataatgaaatgtctatggagaga 435
 DB 468 ggtctggagagccacagggcaaggatgagagagataatggaaggattggggaggagg 527
 OY 436 tgagagagatcccgaggagaaagcttagggagctacagctgagaaattgtcagcatcctta 495

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 11, 2002, 14:17:05 ; Search time 1319.57 Seconds
(Without alignments)
8751.355 Million cell updates/sec

Title: US-09-327-750d-28
Perfect score: 700
Sequence: 1 acagagctctgcccagcagc.....ggggcttctgttcagctga 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8246589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: GenBank

- 1: gb:ba.*
- 2: gb:htg.*
- 3: gb:in.*
- 4: gb:om.*
- 5: gb:ov.*
- 6: gb:pat.*
- 7: gb:ph.*
- 8: gb:pl.*
- 9: gb:pr.*
- 10: gb:ro.*
- 11: gb:sts.*
- 12: gb:sv.*
- 13: gb:un.*
- 14: gb:va.*
- 15: gb:ba.*
- 16: em:fun.*
- 17: em:hum.*
- 18: em:in.*
- 19: em:om.*
- 20: em:or.*
- 21: em:ov.*
- 22: em:pat.*
- 23: em:ph.*
- 24: em:pi.*
- 25: em:sta.*
- 26: em:sv.*
- 27: em:un.*
- 28: em:va.*
- 29: em:htgo_hum.*
- 30: em:htgo_inv.*
- 31: em:htgo_inv.*
- 32: em:htgo_inv.*
- 33: em:htgo_inv.*
- 34: em:htgo_inv.*
- 35: em:htgo_inv.*
- 36: em:htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	692	98.2	700	10 AF187066	AF187066 Mus muscu
2	357.8	50.8	519	10 AF187065	AF187065 Mus muscu
3	337.8	48.3	793	9 BC003190	BC003190 Homo sapi
4	307.4	43.9	891	9 AF187064	AF187064 Homo sapi
5	307.4	43.9	891	9 AF187064	AF187064 Homo sapi
6	240.6	34.4	421	11 G72708	G72708 MARC 4353-4
7	240.6	34.4	421	11 G72708	G72708 MARC 4353-4
8	154.6	22.1	98274	9 AF097439	AF097439 Mus muscu
9	154.6	22.1	98274	9 AF097439	AF097439 Mus muscu
10	101.4	14.5	477	11 G35294	G35294 Human DNA s
11	101.4	14.5	477	11 G35294	G35294 Human DNA s
12	80.2	11.5	1229	9 AK000959	AK000959 Homo sapi
13	80.2	11.5	1229	9 AK000959	AK000959 Homo sapi
14	79.2	11.1	6948	9 HS035649	HS035649 Human DNA
15	67.6	9.7	99	10 M50808	M50808 Homo sapi
16	67.6	9.7	99	10 M50808	M50808 Homo sapi
17	59.8	8.5	7218	6 I66494	I66494 Sequence 14
18	58	8.3	642	9 AF183416	AF183416 Homo sapi
19	58	8.3	828	9 AF220189	AF220189 Homo sapi
20	58	8.3	35714	9 HS198P4	HS198P4 Human DNA
21	57.2	8.2	43952	9 HSU105G4	HSU105G4 Human DNA s
22	56.4	8.1	791	9 AF237783	AF237783 Homo sapi
23	55.8	8.0	744	9 AF251053	AF251053 Homo sapi
24	55.8	8.0	898	6 AX078272	AX078272 Sequence
25	55.8	8.0	31321	9 HS08081	HS08081 Human DNA s
26	55.8	8.0	41029	2 HS08081	HS08081 Human DNA s
27	54.2	7.7	4058	11 G245418	G245418 Human DNA
28	54.2	7.7	4058	11 G245418	G245418 Human DNA
29	54	7.7	835	10 AF097438	AF097438 Mus muscu
30	54	7.7	835	10 AF097438	AF097438 Mus muscu
31	54	7.7	2269	10 AF051347	AF051347 Mus muscu
32	49.2	7.0	185306	2 AC013570	AC013570 Homo sapi
33	48.8	7.0	37224	3 AC007061	AC007061 Leishmani
34	48.2	6.9	36823	9 HSU209G1	HSU209G1 Human DNA s
35	46	6.6	14577	2 AL357046	AL357046 Homo sapi
36	46	6.6	145959	2 AC009282	AC009282 Homo sapi
37	46	6.6	185257	9 AL355305	AL355305 Human DNA
38	44.8	6.4	190544	9 AC005768	AC005768 Homo sapi
39	44.2	6.3	110000	2 LMP1CHR15_5	Continuation (6 of
40	44.2	6.3	172200	2 AC022626	AC022626 Homo sapi
41	43.2	6.2	210211	3 AC070038	AC070038 Dictyosce
42	42.8	6.1	1680	2 P8KCSB	P8KCSB Homo sapi
43	42.6	6.1	164025	2 AC022562	AC022562 Homo sapi
44	42.6	6.1	164025	2 AC022562	AC022562 Homo sapi
45	42.6	6.1	166922	2 AL512288	AL512288 Homo sapi

ALIGNMENTS

RESULT 1	AF187066	700 bp	RNA	11-JUN-2000
DEFINITION	Mus musculus p75NTR-associated cell death executor (Nade) mRNA,			
ACCESSION	AF187066			
VERSION	AF187066.1	GI:8452897		
KEYWORDS	house mouse,			
SOURCE	Mus musculus.			
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
AUTHORS	Wu, J., Hoshino, T., Shoji-Hoshino, S., Kimura, M. T., Nadano, D.,			
TITLE	Signal transduction mediated by the common neurotrophin receptor			
JOURNAL	J. Biol. Chem. 275 (23), 17566-17570 (2000)			
MEDLINE	20298829			

QY	380	gatgaatccgggtttgggtggagatgagatcatatgaggaatcttcatcgaggagatcgag	439
DB	459	TATTGAGCACAA-----TGAAGCGAGAGATGATGTAGAAAGGTTTGTGAGCGAGATGAT	512
QY	440	agagatccggggaagaactctgagagctacagctgagaaatctgctacgcatctctatcggg	499
DB	513	GGMAATCAAGAAAGAAAGACTATGGAGACAGCATAGATGAGCACTATATGCGCTTCCAAA----	568
QY	500	ggagctgtctaaccaaccgcatcacatgatgatgaatctgcttactgctgactgactgcgtc	559
DB	569	-----CTCTGACCTGACAGCATATGATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT	620
QY	560	atccccctcggagatcactgactgactcccgctgagcctctccctccgacttcttc	618
DB	621	AGTTTTCGCTGAGGTTAAATGTGGAACACTGCTTTACAGCTGTATATTTTGTGATTTACT	679
RESULT 14			
H5635G19			
LOCUS		DNA	23-NOV-1999
DEFINITION		Human DNA sequence from clone 635G19 on chromosome 11, 6700 bp. 3	
ACCESSION		H5635G19	
VERSION		1	
KEYWORDS		ESTs and GSSs, complete sequence.	
SOURCE		AL013494	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		Direct Submission	
JOURNAL		Submitted (02-JUN-1999) Sanger Centre, Hinxton, Cambridgeshire,	
COMMENT		CB10 15A, UK. E-mail enquiries: humquerry@sanger.ac.uk	
		Requests: clonerequests@sanger.ac.uk	
		On May 11, 1999 this sequence version replaced gi:4587749.	
		During differences assembly data is compared from overlapping clones.	
		The following abbreviations are used to associate primary accession	
		numbers given in the feature table with their source databases:	
		Emi: EMBL; Sw: SWISSPROT; Tr: TREMBL	
		This sequence has been finished according to sequence map criteria	
		as follows. An attempt is made to resolve all sequencing problems,	
		such as compressions and repeats, but not necessarily within known	
		annotated human repeat sequence elements (e.g. Alu). Where the	
		sequence is ambiguous, there is an annotation using the 'unsure'	
		feature key.	
		This sequence was generated from part of bacterial clone contigs of	
		H5635G19. The sequence was constructed by the Sanger Centre Chromosome X	
		Mapping Group. Further information can be found at	
		http://www.sanger.ac.uk/HGP/ChrX	
		635G19 is from the library RP014 constructed at the Roswell Park	
		Cancer Institute by the group of Plecter de Jong. For further	
		details see http://bacpac.med.buffalo.edu/VECTOR: pcYPAC2	
		IMPORTANT: This sequence is not the entire insert of clone 635G19.	
		It may be shorter because we only sequence overlapping sections	
		once, or longer because we arrange for a small overlap between	
		neighbouring submissions.	
		The true left end of clone U63A (281014) is at 65849 in this	
		sequence. The true right end of clone U10D3 (283997) is at 100 in	
		this sequence.	
FEATURES		1..65848	
source		/organism='Homo sapiens'	
		/db_xref='taxon:9606'	

```

PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
  Template: 25 ng
            each 1 uM
  Primers:  each 200 uM
  dNTPs:     each 200 uM
  Taq Polymerase: 0.05 units/ul
  Total Vol: 10 ul

  Buffer: MgCl2: 2.5 mM
         KCl: 50 mM
         Tris-HCl: 20 mM
         pH: 8.3

Prepared with primer pairs derived from M38188 -- Unigene.

FEATURES
Source          Location/Qualifiers
STS             1..244
                organelle="Mitochondrion"
                organism="Homo sapiens"
primer_bind     84..103
primer_bind     199..221
BASE COUNT      57 a 50 c 42 g 95 t

ORIGIN
Query Match    11.18; Score 77.4; DB 11; Length 244;
Best Local Similarity 80.24; Fred. No. 8.6e-10;
Matches 105; Conservative 0; Mismatches 21; Indels 5; Gaps 1;

Qy 567 cctggagatcatcactgacccgcgctgacgcttcctgcgctttctcgcacatgcc 626
Db 15 CAGGAGATTAATACTGTGATCCCGCTGGTTCTTTCTTCCTTCGATTTTCTTAATATGCC 74

Qy 627 tttaatgacccgtttgtgtgacccgtgtttatttcacatgccatgacgagtgaggcct 686
Db 75 TTACTAGTCCGTTTGCTGTGAACCACTATGTTATT-----CCATGTGTCAAGTGGGTCT 129

Qy 687 tgtgttgccag 697
Db 130 TGTGTGTGCCAG 140

Search completed: March 11, 2002, 16:33:03
Job time: 815f sec

```

Result	Score	Query	Length	DB	ID	Description
1	687.4	98.2	809	12	AK004531	AK004531 Mus muscu
2	687.4	98.2	845	12	AK010500	AK010500 Mus muscu
3	637.9	91.0	787	11	BF608455	BF608455 MY1_00139
4	627.6	89.7	642	10	A1118980	A1118980 uc93c08.Y
5	606.8	86.7	628	10	A4476968	AM76468 uc976002.Y
6	591.8	84.4	599	11	BG088461	BG088461 H3153007.Y
7	590.4	84.3	622	10	A272675	BD72375 VB26907.Y
8	589.8	82.6	575	10	B1407164	B1407164 G0919806
9	578.8	82.6	877	11	E3434877	E3434877 uc90310.Y
10	571.6	81.7	895	15	BE334877	BE334877 uc90310.Y
11	566.6	80.9	993	11	BF178306	BF178306 60380845.Y
12	535.6	76.5	542	10	BE334866	BE334866 uc90411.Y


```
|||||
Db 1 CCCGGGGGCTGGCAGCGGGCCCGAGCGGAGCGGACAGATTGACTGGAAGCCGAGAGTC 60
QY 97 cagggcgagcggggaattgacagggagactacgcccgaagggtataggcccaagaatagcaa 156
|||||
Db 61 CAGCGCGGACCGGGAAATTGACNAGGAGACTACGCCCGNAGGATAGGCCCCNAGATAGCA 120
QY 157 cagggaaacaatctcatcgccaatctccaccaggaacagaaagacttgagcagc 216
|||||
Db 121 CAGGAACAATAATCTCATATGCCCAATGTCCACGAGGAACGAAGATGGAGCAGC 180
QY 217 ccctgcagaatggacaggaacacccgcccctgtggagaggtgagggccaccagcctgctg 276
|||||
Db 181 CCCTGCAGATGGACAGGAAGACCGCCCTGTGGGAGAGGTGAGGGCCACCCAGCCTGCTG 240
QY 277 caaiaacaacaacaacaacaacaacaacaacaacaacaacaacaacccagagggcaggctc 336
|||||
Db 241 CAACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 300
QY 337 gccgaacctgcccctaactctccggtggggcattcccaacagcgagatgacgggttgg 396
|||||
Db 301 GCCGACTTGGCCCTAACTTCCGATGGGCCATTCCCACAGCGAGATGAATGACGGGTTGG 360
QY 397 gtggagctggagatgatggaaatgttcctatggagagatgagagagatccggagaaagc 456
|||||
Db 361 GTGGAGATGGAGATCATGGAAATGTTTCATGGAGGAGATGAGAGATCCGGAGAAAGC 420
QY 457 ttagggactacagctgagaattgctacgataccttatggggagctgtctaaccacc 516
|||||
Db 421 TTAGGAGCTACAGCTGAGAAATTTGCTACGCCATCTTATGGGGAGCTGTCTAACACCC 480
QY 517 acgatcacatgatgaattctgcttctgcttctgcttctgcttctgcttctgcttctgct 576
|||||
Db 481 ACGATCACCATGATGATCTGCTTATGCTTACTTGGTTCGCTCATTCG-CCCTGAGATCC 539
QY 577 atactgtactcc 589
|||||
Db 540 ATACTGTGACTCC 552
```

Search completed: March 11, 2002, 16:09:50
Job time: 10885 sec

	Query Match	Best Local Similarity	5.3%	Score 36.8	DB 4	Length 289	
	Matches 15	Conservative	99	Mismatches 118	Indels	0	Gaps 0
98	agcgcgagcggaattgacgaggaagctacgcgcgaggaatagagccaggaatagcaac	157					
1	rgtggatcctatctatctgcgcgaattgcctacgcgaggaagctgagcgagcc	60					
158	caggagaaacaaatctcatctgcgcgaattgcctacgcgaggaagctgagcgagcc	217					
61	rfsnrnsnr	120					
218	ctcgcaatagacaggaacacgcgcctctgtggagggagttagggccacagctgtgc	277					
121	rfsnrnsnr	180					
278	aac	329					
181	rfsnrnsnr	232					

```

RESULT      5
US-09-244-796-17
: Sequence 17, Application US/09244796
US-09-244-796-17
: Patent NO. 6281344
: GENERAL INFORMATION:
: APPLICANT: Szoatak, Jack W.
: APPLICANT: Roberts, Richard W.
: INVENTOR: Szoatak, Jack W.
: TITLE OF INVENTION: THE SELECTION OF PROTEINS USING RNA-PROTEIN
: TITLE OF INVENTION: FUSIONS
: FILE REFERENCE: 00786/310007
: CURRENT APPLICATION NUMBER: US/09/244.796
: CURRENT FILING DATE: 1999-02-05
: EARLIER APPLICATION NUMBER: 60/035,963
: EARLIER FILING DATE: 1997-01-27
: EARLIER APPLICATION NUMBER: 60/064,491
: EARLIER FILING DATE: 1997-01-27
: EARLIER APPLICATION NUMBER: 09/007,005
: EARLIER FILING DATE: 1998-01-14
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 289
: TYPE: RNA
: FEATURE: SM
: FEATURE: OTHER INFORMATION: Translation template
: OTHER INFORMATION: Translation template

```

```

// NUMBER OF SEQ ID NOS: 33
// SOFTWARE: FastSeq for Windows Version 4.0
// SEQ ID NO 17
// LENGTH: 289
// TYPE: RNA
// ORGANISM: Artificial Sequence
// FEATURE:
// OTHER INFORMATION: Translation template

```

```

1 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
2 TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
3 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
4 TITLE OF INVENTION: SPECIES INFECTIONS
5 NUMBER OF SEQUENCES: 30
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: PETERS, VERNY, JONES & BIKSA
8 STREET: 385 Sherman Avenue, Suite 6
9 CITY: Palo Alto
10 STATE: CA
11 COUNTRY: USA
12 ZIP: 94306-1840
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: Patentin Release #1.0, Version #1.30
18 CURRENT APPLICATION DATA: /08/928, 361B
19 FILING DATE: 12-SEP-1997
20 CLASSIFICATION:
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 60/026,062
23 FILING DATE: 13-SEP-1996
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Verny, Hana
26 REGISTRATION NUMBER: 30,518
27 REFERENCE/DOC#T NUMBER: 480,76-1(HV)
28 TELEPHONE: 650-324-1677
29 TELEFAX: 650-324-1678
30 INFORMATION FOR SEQ ID NO: 3:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 5318 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: double
35 TOPOLOGY: linear
36 MOLECULE TYPE: DNA (genomic)
37 US-08-928-361B-3
38
39 Query Match 5.0%; Score 34.8; DB 3; Length 5318;
40 Best Local Similarity 84.8%; Pred. No. 1.3;
41 Matches 39; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
42
43 QY 279 aaacaaacacacacacacacacacacacacacacacacacacacacacacgaa 324
44 ||| ||||| ||||||| ||||||| ||||||| ||||||| ||||| |||
45 DB 873 AACTACACACACACACACACACACACACACACACACACACACCACTACCAAGNA 918
46 ||| ||||| ||||||| ||||||| ||||||| ||||||| ||||| |||
47
48 RESULT 11
49 US-08-692-922-1/G
50 Sequence 1, Application US/08692922
51 Patent No. 6277592
52 GENERAL INFORMATION:
53 APPLICANT: Bidwell, Christopher A.
54 APPLICANT: Spurluck, Michael E.
55 TITLE OF INVENTION: PORCINE LEFTIN PROTEIN, NUCLEIC ACID
56 TITLE OF INVENTION: SEQUENCES CODING THEREFOR AND USES THEREOF
57 NUMBER OF SEQUENCES: 30
58 CORRESPONDENCE ADDRESS:
59 ADDRESSEE: WYTE HIRSCHBOECK DUDEK S.C.
60 STREET: Suite 2100 111 East Wisconsin Avenue
61 CITY: Milwaukee
62 STATE: Wisconsin
63 COUNTRY: USA
64 ZIP: 53202
65 COMPUTER READABLE FORM:
66 MEDIUM TYPE: Floppy disk
67 COMPUTER: IBM PC compatible
68 OPERATING SYSTEM: PC-DOS/MS-DOS
69 SOFTWARE: Patentin Release #1.0, Version #1.30
70 CURRENT APPLICATION DATA:

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sv model

Run on: March 11, 2002, 16:36:39 ; Search time 144 Seconds
(Without alignments)
5304.700 Million cell updates/sec

Title: US-09-327-750d-29
Perfect score: 891
Sequence: 1 accaccatccccctctctat.....ataaagcaattataaagc 891

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs. 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889.4	99.8	891	22	AAF23529
2	564.4	53.3	917	21	AAF21748
3	314.3	38.5	524	21	AAO01029
4	298.6	33.5	495	21	AAO01004
5	273.8	30.7	426	22	AAH2586
6	273.8	30.7	426	22	AAH2586
7	160.8	18.0	169	16	AAAT1034
8	150.8	16.9	187	21	AAH7147
9	141.2	15.8	865	22	AAH03517
10	141.2	15.8	1229	22	AAH13750
11	141.2	15.8	1364	22	AAH85548

12	108.2	12.1	792	22	AAH45143
13	104	11.7	988	22	AAF59611
14	102.4	11.5	767	22	AAF73810
15	102.4	11.5	858	22	AAH2586
16	91.4	10.3	858	22	AAH2586
17	91.4	10.3	692	21	AAO10889
18	89.8	10.1	562	21	AAO3880
19	89.8	10.1	598	21	AAO3879
20	81	9.1	936	22	AAF58252
21	81	9.1	936	22	AAF58254
22	81	9.1	936	22	AAF58257
23	81	9.1	936	22	AAF58259
24	81	9.1	936	22	AAF58252
25	74.9	8.7	1101	21	AAO11003
26	69.5	7.7	1101	21	AAO11003
27	67.4	7.6	936	22	AAF58252
28	67.4	7.6	936	22	AAF58254
29	67.4	7.6	936	22	AAF58257
30	67.4	7.6	936	22	AAF58259
31	67.4	7.6	936	22	AAF58252
32	67.4	7.6	936	22	AAF58254
33	67.4	7.6	936	22	AAF58257
34	66.4	7.5	128	21	AAO12100
35	64.8	7.3	397	22	AAI14092
36	64.8	7.3	397	22	AAI14092
37	64.8	7.3	397	22	AAI14092
38	58.6	6.6	1250	22	AAI57960
39	58.6	6.6	1250	22	AAI57960
40	58.6	6.6	1551	22	AAI57960
41	56.4	6.3	356	21	AAH30950
42	55.4	6.2	1309	21	AAH30950
43	52.8	5.9	285	22	AAI21436
44	52.8	5.9	285	22	AAI21436
45	52.8	5.9	285	22	AAI6725
			285	22	AAI07131

ALIGNMENTS

RESULT	1
ID	AAF23529 standard; DNA; 891 BP.
AC	AAF23529;
DT	22-MAR-2001 (first entry)
DE	Human NADE DNA.
KW	Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW	neurogenetic disease; NF-kappaB; ds.
XX	Homo sapiens.
XX	MO200075278-A2.
PD	14-DEC-2000.
PF	07-JUN-2000; 2000MO-US15621.
PR	07-JUN-1999; 99US-0327750.
XX	(UYCO) UNIV COLUMBIA NEW YORK.
XX	Sato T.
XX	WPI; 2001-061707/07.
PT	New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT	the gene encoding NADE, useful for modulating the activity of p75NTR
PT	and for detecting neurodegenerative diseases -
XX	Disclosure; Fig 1; 134pp; English.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that are used in diagnostic, forensic, gene therapy and chromosome mapping procedures -

Claim 1: SEQ ID 1003: 71bp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or poly(A)⁺ RNAs derived from different tissues. EST sequences usually correspond to the 5' ends of cDNAs and genomic DNAs. Such ESTs are not often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

Sequence 532 BP: 151 A: 121 C: 163 G: 91 T: 6 Other: 50

368	catcacccataggcagacacaaagcgggagatgggacgagatggagatacgggaataact	447
369		
370		
371		
372		
373		
374		
375		
376		
377		
378		
379		
380		
381		
382		
383		
384		
385		
386		
387		
388		
389		
390		
391	ggatgcctcatggggagagctctcagacacacagacacacacacga	625
392		
393		
394		
395		
396		
397		
398	ggatgcctcatggggagagctctcagacacacacacacacga	632
399		
400		
401		
402		
403		
404		
405		
406		
407		
408		
409		
410		
411		
412		
413		
414		
415		
416		
417		
418		
419		
420		
421		
422		
423		
424		
425		
426		
427		
428		
429		
430		
431		
432		
433		
434		
435		
436		
437		
438		
439		
440		
441		
442		
443		
444		
445		
446		
447		
448		
449		
450		
451		
452		
453		
454		
455		
456		
457		
458		
459		
460		
461		
462		
463		
464		
465		
466		
467		
468		
469		
470		
471		
472		
473		
474		
475		
476		
477		
478		
479		
480		
481		
482		
483		
484		
485		
486		
487		
488		
489		
490		
491		
492		
493		
494		
495		
496		
497		
498		
499		
500		
501		
502		
503		
504		
505		
506		
507		
508		
509		
510		
511		
512		
513		
514		
515		
516		
517		
518		
519		
520		
521	catcgaggagatcggagaaatcagagaaaaactcaggagactcaggttcgggaattgct	580
522		
523		
524		
525		
526		

db	488	gcgtatccttatgggggaktctctctaatcaccatgacctatga	532
	RESULT		
	4		
	AF23528		
	1	AA23528 standard; DNA; 700 BP.	
	XX		
	XX	AA23528;	
	XX		
	XX	22-MAR-2001 (first entry)	
	XX		
	XX	Mouse NADE DNA.	
	XX		
	XX	Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;	
	XX	neurogenetic disease; NF-kappaB, ds.	
	XX		
	XX	Mus sp.	
	XX		
	XX	WO200075278-A2.	
	XX		
	XX	14-DEC-2000.	
	XX		
	XX		


```
CC specification.
XX
SQ Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

Query Match      11.5%; Score 102.4; DB 22; Length 862;
Best Local Similarity 55.3%; Percent Ids 20;
Matches 345; Conservative 0; Mismatches 241; Indels 38; Gaps 6;

QY 283 aanaacacacagaaanaaaatctcatcgtgcgaatattccaccaggaagaaagaga 342
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 aagaggaacgagcgttaacacatctcatcgtggaaaatgtcaaccaggaanaatgaaa 295
QY 343 tggagcagcctatgcagaatggaggaaga-----ccgccctttggaggag 390
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
296 aagatgaaaaggagcaggtgtctaatgaaggagggcccttggccctacctttgaatgta 355
QY 391 gtgaagggccacccgctgcaggaas--tcagagggaacagctgcagactgccccta 447
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
356 gtgaactactgtgctcagagggaacagcaggcgttcgcgcttaagcagccatctgc 415
QY 448 atttctgagtgccatacccaatagcagcatcaatgatggagtggtggagatggagatg 507
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
416 agtatagatggagacataatgcataggcttgagagccacagcgaaggatgagagagaga 475
QY 508 atatggaaatattcatggaggagatgagagaatcagaanaaacttagggagctgcagt 567
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
476 atatggaaaggatcgggagagggtgagacagctgatggaanaagctgaggaanaagcagt 535
QY 568 tgaagaaattctctgcgtatcctcttatggggggagctctctaataccatgacatcatgatg 627
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
536 tgaatcagtttgcgggcagtcagcactgacccccc---tcactatgacatcatcagtg 592
QY 628 aatttgccttatgcttgcactctccatttatatagagattaatactgtattccgcg 687
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
593 agtttgccttatgcttgcactctccatttatggtttccctgaagttaataggagagaccctgc 652
QY 688 tgtttcttttttcttgcatttctccaatagccttactgtatccgtttgctgaaaccc 747
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
653 ttcctaaacttaca-----cattggtgggtacctttgtcgtaaaag 694
QY 748 tatgtattctcatgtctcaatggtggtggtggtggtggtggtggttctattgaagattgct 807
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
695 ttgtgatgttactattcttctgtggtggtcctctattaccagcttctaattgaatgtttt 754
QY 808 ttg-cactcagtgtaagtctctgcagcagtag-ttccaccctattgcatggaanaattt 865
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
755 ttgaccacagtttgtaagtttctgacgacgagaggtttaccctattgcatggaagagtc 814
QY 866 aaagccataaagcaattcaaaa 889
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
815 tcattatatttgaagtttaata 838
```

Search completed: March 11, 2002, 16:36:43
Job time: 3232 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 11, 2002, 16:33:03 ; Search time 1319.57 Seconds
11139-225 Million cell updates/sec

Title: US-09-327-750d-29
Perfect score: 891
Sequence: 1 accaccatccccctctat.....ataaagcaatttaaaagc 891
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824859755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.com.*
5: gb.con.*
6: gb.ph.*
7: gb.pl.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.un.*
12: gb.vi.*
13: gb.y.*
14: gb.yi.*
15: en.ba.*
16: en.in.*
17: en.ph.*
18: en.pl.*
19: en.pr.*
20: en.ro.*
21: en.sts.*
22: en.un.*
23: en.vi.*
24: en.y.*
25: en.yi.*
26: en.ba.*
27: en.con.*
28: en.htg.*
29: en.in.*
30: en.ph.*
31: en.pl.*
32: en.pr.*
33: en.ro.*
34: en.sts.*
35: en.un.*
36: en.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	899.4	99.8	891	9 HUMOGC	M38188 Homo sapi
2	899.4	99.8	891	9 AF187064	AF187064 Homo sapi
3	609.4	68.4	793	9 BC003190	BC003190 Homo sapi
4	363	40.7	854	10 AF097440	AF097440 Mus muscu
5	351.2	39.6	700	10 AF187065	AF187065 Mus muscu
6	341.2	37.5	519	10 AF187065	AF187065 Rattus no
7	245.4	27.5	244	11 G13550	G13550 human STS S
8	245.4	27.5	244	11 G13550	G13550 human STS S
9	185.2	20.8	45678	9 HSU351F8	HSU351F8 Human DNA
10	185.2	20.8	45678	9 HSU351F8	HSU351F8 Human DNA
11	138.6	17.8	69648	9 HSU351F8	HSU351F8 Human DNA
12	141.2	15.8	1229	9 AK000959	AK000959 Homo sapi
13	141.2	15.8	1364	6 AX100231	AX100231 Sequence
14	141	15.8	43952	9 HSU10504	HSU10504 Human DNA
15	132.4	14.9	477	11 G35294	G35294 human STS S
16	122.2	13.7	35214	9 HSU3984	HSU3984 Human DNA
17	117.2	13.2	13457	9 HSU3984	HSU3984 Human DNA
18	117.2	13.2	13457	9 HSU3984	HSU3984 Human DNA
19	117.2	13.2	13457	9 HSU3984	HSU3984 Human DNA
20	110.8	12.4	31321	9 HSU87088	HSU87088 Human DNA
21	110.8	12.4	41029	2 HSU8081	HSU8081 Homo sapi
22	109.8	12.3	642	9 AF183416	AF183416 Homo sapi
23	109.8	12.3	828	9 AF220189	AF220189 Homo sapi
24	109.8	12.3	40584	9 AL133348	AL133348 Homo sapi
25	108.2	12.1	791	9 AF237783	AF237783 Homo sapi
26	104	11.7	898	6 AX078272	AX078272 Sequence
27	102.4	11.5	744	9 AF251053	AF251053 Homo sapi
28	77	8.6	43952	9 HSU10504	HSU10504 Human DNA
29	74.6	8.4	1163	9 BC005988	BC005988 Homo sapi
30	74.6	8.4	1163	9 BC005988	BC005988 Homo sapi
31	69.4	7.0	728	10 AF097439	AF097439 Mus muscu
32	59.6	6.7	40822	9 HSU17758	HSU17758 Human DNA
33	59.4	6.7	13872	2 AC027560	AC027560 Homo sapi
34	59.4	6.7	176039	2 AC073065	AC073065 Homo sapi
35	59.4	6.6	1208	9 AK026349	AK026349 Homo sapi
36	58.6	6.6	91480	9 HSDJ823F3	AL079333 Human DNA
37	58.6	6.6	156275	2 AL513285	AL513285 Homo sapi
38	58.4	6.4	196367	9 AC020740	AC020740 Homo sapi
39	57	6.4	2269	10 AF097437	AF097437 Mus muscu
40	56.6	6.4	18896	2 AC037486	AC037486 Homo sapi
41	56.6	6.3	127606	2 AL556298	AL556298 Homo sapi
42	56.2	6.3	19013	2 AC037486	AC037486 Homo sapi
43	56.2	6.3	19013	2 AC037486	AC037486 Homo sapi
44	56.2	6.3	19013	2 AC037486	AC037486 Homo sapi
45	56.2	6.3	19013	2 AC037486	AL079303 Human Chr

ALIGNMENTS

RESULT 1	HUMOGC	891 bp	mRNA	PRI : 07/04/1995
LOCUS	Human unknown protein from clone PHGR74			complete cds.
DEFINITION	Human unknown protein from clone PHGR74			
ACCESSION	M38188.1			
VERSION	M38188.1			
KEYWORDS	GI:189378			
SOURCE	Human ovular granulosa cell line, cDNA to mRNA.			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 891)			
AUTHORS	Zimmer, M. and Scheit, K.H.			
TITLE	Characterization of three abundant mRNAs from human ovarian granulosa cells.			
JOURNAL	Journal of Biol. Chem.			
MEDLINE	9102550			
COMMENT	Draft entry and computer-readable sequence for [DNA 9, 479-485 (1990)] kindly submitted			

Db 1 ACCCCATCCCACTCTATACCGGCTCTCCATTTTGGTGCCTGCAACGCTCTGGGAAG 60
Qy 61 atccccgggaacgaaaaatggtggatttgggggaagggggaaggggaagctggga 120
Db 61 AATCCCGGGAACGAAAAATGCGTGTGGGGGAGGAGGTAAAGGAGAAAGCTCGA 120
Qy 121 gggaggggcttaattgagagcccgtagagagcgcgggaacttcaaggtggaaaaa 180
Db 121 GGGAGGGCTTAAATGAGGCCCGTAGAGAGCGCGGGAATCTCTAGGTGGGAAAA 180
Qy 181 acgaaatcaaaaccccttggatcatcagggcctcgaatcctgctgagcagcaccagc 240
Db 181 ACGAAATAAAAATCTTTGATATCAGGGCTGTAATCTGCTGCTGAGAGCAGCAAGC 240
Qy 241 attcagctctctcctggcttctgcttactctgtctaaagaagaaaaaacacccgaaaaa 300
Db 241 ATTCACT 300
Qy 301 aaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 360
Db 301 AAATCT 360
Qy 361 atgag 420
Db 361 ATGGAG 420
Qy 421 gggag 480
Db 421 GGGAG 480
Qy 481 atgag 540
Db 481 ATGATGGAG 540
Qy 541 tcaag 600
Db 541 TCAGAAG 600
Qy 601 tctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 660
Db 601 TCTCTATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 661 tcaag 720
Db 661 TCATGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 721 ctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 780
Db 721 CTTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 781 ttgcag 840
Db 781 TTGCAG 840
Qy 841 ttccacccattgag 891
Db 841 TTCACCCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 891

RESULT 3
BC003190 793 bp mRNA PRI 12-JUL-2001
LOCUS Homo sapiens, p75NTR-associated cell death executor; ovarian
DEFINITION p75NTR-associated cell death executor; ovarian
organism: Homo sapiens
ACCESSION BC003190.1 GI:13112030
VERSION BC003190.1
KEYWORDS MGC
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 793)
Srausberg, R.
Direct Submission
Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC Mail Room
Email: mgc@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fyell, Erin Garland, Fan Guin,
Leticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo
San Lee, Victor Ling, Carrie Matheson, Kandice McCreary, Steven
Schmid, Duane Smalhus, Michael Smith, Patricia Smith, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone Distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 6 Row: k Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7657043.
Location/Qualifiers
i. 793;am="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:802 IMAGE:3357965"
/tissue_type="eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
196..531
/codon_start=1
/product="p75NTR-associated cell death executor; ovarian
p75NTR-associated cell death executor (13kb)"
/db_xref="GI:13112031"
/translation="MANHOENEMEPONGNEDEPPLGGGGHOPAGNRRGOARRLA
PNFRMAIPNRQINDMGDDMDIFNEMREIRKRLRLNLRNLRNLRNLRNLRNLRNLRNLRN
HHDFELMP"
BASE COUNT 225 a 170 c 205 g 193 t
ORIGIN
Query Match 58.4%; Score 609.4; DB 9; Length 793;
Best Local Similarity 99.8%; Read Nucleotide 15;
Matches 610; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 281 agaaaaaacacacgagaaaaaaatctctctgagcaaatattccacagaggaacgagga 340
Db 165 AGAAAAACAACACAGAAAAAAAATCTCATGTGCAAAATATTCCACAGGAAACGAGA 224
Qy 341 gatggagagcctcagcaaatgagagagagagagagagagagagagagagagagagga 400
Db 225 GATGGAGAGCCTATGAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
Qy 401 cgaagctcgaaggaatcag 460
Db 285 CCACCTCGAGGAATCGAGGGGCAAGCGCTCCCGACCTGCCCCCTAATTTTCATGG 344
Qy 461 cataccatagcagagatcatgagtgagtgagtgagtgagtgagtgagtgagtgag 520
Db 345 CATACCAATAGCAGATCAATGATGGATGGATGGATGGATGGATGGATGGATGGAT 404

Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers in the GenBank database with their source databases:
 BAC: BAC library
 EST: EST database
 FISH: FISH database
 on the MRPPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/Wormpep CTA-71487 is from the human BAC library described in U-J. Kim et al. (1996) Genomics 34, 213-218.
 VECTOR: pBeloBAC11
 This sequence is the entire insert of clone CTA-71487.

FEATURES

source

1. 38274
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="22"
 /map="q12.2-q13.2"
 /clone="CTA-71487"
 /clone.lib="CIT578SK-A2"
 109..136
 /note="14 copies 2 mer ac 100 conserved"
 repeat_region
 183..258
 /note="LMCS repeat: matches 7840..7913 of consensus"
 repeat_region
 1177..1446
 /note="match: GSS: Em:AQ373724"
 misc_feature
 1301..11356
 /note="11356 repeat: matches 2693..2750 of consensus"
 repeat_region
 2558..5036
 /note="LMOD1 repeat: matches 3673..6170 of consensus"
 repeat_region
 5036..5486
 /note="LM8C repeat: matches 2052..2207 of consensus"
 repeat_region
 5496..6885
 /note="LM8C repeat: matches 2052..2207 of consensus"
 repeat_region
 6896..7185
 /note="LMD repeat: matches 1..1538 of consensus"
 repeat_region
 complement(7187..7611)
 /note="L2 repeat: matches 2356..2710 of consensus"
 misc_feature
 7369..7394
 /note="match: GSS: Em:AQ566001"
 repeat_region
 complement(7395..8093)
 /note="LMR repeat: matches 34..256 of consensus"
 misc_feature
 complement(7395..8093)
 /note="match: GSS: Em:AQ071449"
 repeat_region
 7304..7578
 /note="MIR repeat: matches 20..95 of consensus"
 repeat_region
 8086..8300
 /note="MIR repeat: matches 8..255 of consensus"
 misc_feature
 8154..8565
 /note="match: GSS: Em:AQ059599"
 repeat_region
 8549..8864
 /note="AluX repeat: matches 1..280 of consensus"
 misc_feature
 complement(8865..9290)
 /note="match: GSS: Em:AQ143214"
 repeat_region
 9235..9361
 /note="L2 repeat: matches 2619..2749 of consensus"
 misc_feature
 9378..9890
 /note="match: GSS: Em:AQ284832"
 repeat_region
 10135..10428
 /note="AluX repeat: matches 1..294 of consensus"
 repeat_region
 10429..10563
 /note="AluX repeat: matches 1..135 of consensus"
 repeat_region
 10811..10868
 /note="L2 repeat: matches 2648..2705 of consensus"
 repeat_region
 10819..11064
 /note="MIR repeat: matches 2..255 of consensus"
 repeat_region
 11112..11132
 /note="MIR repeat: matches 11..255 of consensus"
 repeat_region
 11869..12012
 /note="MIR repeat: matches 20..167 of consensus"

12925..13223
 /note="AluY repeat: matches 1..300 of consensus"
 repeat_region
 13290..13495
 /note="MER38A repeat: matches 9..214 of consensus"
 misc_feature
 13338..13833
 /note="match: GSS: Em:AQ568864"
 repeat_region
 13688..13749
 /note="MIR repeat: matches 77..139 of consensus"
 repeat_region
 13840..13966
 /note="MIR repeat: matches 22..147 of consensus"
 repeat_region
 14182..14194
 /note="MIR repeat: matches 46..136 of consensus"
 repeat_region
 14925..15058
 /note="MIR repeat: matches 48..191 of consensus"
 repeat_region
 15994..16040
 /note="MIR repeat: matches 102..143 of consensus"
 repeat_region
 16116..16184
 /note="MER39 repeat: matches 3..70 of consensus"
 repeat_region
 16897..17005
 /note="MIR repeat: matches 9..118 of consensus"
 repeat_region
 17618..17700
 /note="MIR repeat: matches 108..192 of consensus"
 repeat_region
 17718..17919
 /note="MIR repeat: matches 73..245 of consensus"
 repeat_region
 18657..18871
 /note="MIR repeat: matches 35..252 of consensus"
 repeat_region
 18927..19230
 /note="AluX repeat: matches 1..301 of consensus"
 repeat_region
 19492..19559
 /note="MIR repeat: matches 2..67 of consensus"
 repeat_region
 19560..19821
 /note="AluY repeat: matches 34..295 of consensus"
 repeat_region
 19822..19959
 /note="LMOD1 repeat: matches 67..212 of consensus"
 repeat_region
 19987..20386
 /note="AluY repeat: matches 1..300 of consensus"
 repeat_region
 20316..20512
 /note="MIR repeat: matches 82..262 of consensus"
 repeat_region
 20513..20741
 /note="LMOD1 repeat: matches 5970..6224 of consensus"
 repeat_region
 20769..21074
 /note="AluSg1 repeat: matches 1..303 of consensus"
 repeat_region
 21085..21207
 /note="LM2 repeat: matches 1776..1898 of consensus"
 repeat_region
 21212..21391
 /note="LM2 repeat: matches 6..178 of consensus"
 repeat_region
 21392..21419
 /note="MIR repeat: matches 3..86 of consensus"
 repeat_region
 22146..22441
 /note="AluX repeat: matches 1..296 of consensus"
 repeat_region
 22526..22610
 /note="LTR16C repeat: matches 267..349 of consensus"
 repeat_region
 22940..23204
 /note="HERV16 repeat: matches 378..650 of consensus"
 repeat_region
 23205..23514
 /note="AluSg1 repeat: matches 1..311 of consensus"
 repeat_region
 23919..24194
 /note="MIR repeat: matches 36..311 of consensus"
 repeat_region
 24456..24590
 /note="LTR16C repeat: matches 253..349 of consensus"
 repeat_region
 24871..24916
 /note="MIR repeat: matches 102..143 of consensus"
 repeat_region
 25378..25555
 /note="L2 repeat: matches 2330..2519 of consensus"
 repeat_region
 25673..25910
 /note="MIR repeat: matches 4..255 of consensus"
 repeat_region
 26722..26798
 /note="L2 repeat: matches 2572..2649 of consensus"
 repeat_region
 27040..27107
 /note="MIR repeat: matches 2489..2500 of consensus"
 repeat_region
 27108..27157
 /note="AluY repeat: matches 41..304 of consensus"
 repeat_region
 27375..27471


```

RESULT 10
LOCUS   HSV351F8      45678 bp      DNA      23-NOV-1999
DEFINITION   Human DNA sequence from cosmid V351F8, between markers DXS366 and
              DXS87 on chromosome X contains ESTs.
ACCESSION   270719
VERSION     1
KEYWORDS    X
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 45678)
AUTHORS    Whiteley M.
TITLE      Direct Submission
JOURNAL    Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1HQ, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT    REPORT: This sequence is the entire insert of clone V351F8. The
            right end of clone V351F8 is at 45678.
            V351F8 is from the human chromosome X-specific cosmid library.
FEATURES             Location/Qualifiers
     source            1..45678
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="X"
                     /clone="GHC-351F8"
                     /clone_lib="SCCV"
     repeat_region     1..76
                     /note="Alu repeat: matches 80. .1 of consensus"
     repeat_region     642..930
                     /note="Alu repeat: matches 1. .308 of consensus"
     repeat_region     1406..5179
                     /note="L1 element fragment"
     repeat_region     5243..5276
                     /note="17 copies of 2 mer 82 & conserved"
     repeat_region     5279..5567
                     /partial
     repeat_region     9775..10369
                     /note="Alu repeat: matches 304. .1 of consensus"
     repeat_region     10369..10520
                     /note="MSTR element fragment"
     repeat_region     10520..11197
                     /note="MSTR element fragment"
     repeat_region     11318..11978
                     /note="MSTR element fragment"
     repeat_region     11715..11846
                     /note="MSTR element fragment"
     repeat_region     12074..12170
                     /note="MSTR element fragment"
     repeat_region     12227..12486
                     /note="MSTR element fragment"
     repeat_region     12437..13502
                     /note="MSTR element fragment"
     repeat_region     12756..13059
                     /note="MSTR element fragment"
     repeat_region     14478..14650
                     /note="Alu repeat: matches 308. .1 of consensus"
     repeat_region     14491..14536
                     /note="MSTR element fragment"
     repeat_region     14715..14831
                     /note="MSTR element fragment"
     repeat_region     14899..15333
                     /note="MSTR element fragment"
     repeat_region     15534..15620
                     /note="MSTR element fragment"
     repeat_region     15566..15619
                     /note="31 copies of 3 mer 85 & conserved"
     repeat_region     16352..16436
                     /note="3 copies of 18 mer 98 & conserved"
     repeat_region     16728..16763
                     /note="L1 element fragment"

```

```

repeat_region     16764..17054
                     /partial
     repeat_region     19420..19449
                     /note="Alu repeat: matches 308. .1 of consensus"
     misc_feature     19420..19449
                     /note="15 copies of 2 mer 87 & conserved"
     repeat_region     20013..20013
                     /note="match: 3' EST N51315 clone 283089"
     repeat_region     20073..20112
                     /note="22 copies of 2 mer 98 & conserved"
     repeat_region     20073..20126
                     /note="10 copies of 4 mer 100 & conserved"
     repeat_region     20073..20126
                     /note="3 copies of 18 mer 87 & conserved"
     repeat_region     23088..23300
                     /note="L1 element fragment"
     repeat_region     23311..23490
                     /note="MSTR element fragment"
     repeat_region     23318..23489
                     /note="MSTR element fragment"
     repeat_region     23523..23489
                     /note="MSTR element fragment"
     repeat_region     23507..24407
                     /note="L1 element fragment"
     repeat_region     24376..24462
                     /note="MSTR element fragment"
     repeat_region     24379..24491
                     /note="MSTR element fragment"
     repeat_region     24688..24738
                     /note="MSTR element fragment"
     repeat_region     24810..25016
                     /note="MSTR element fragment"
     repeat_region     25251..25632
                     /note="L1 element fragment"
     repeat_region     25439..25696
                     /note="MER25 element fragment"
     repeat_region     26284..26343
                     /note="MSTR element fragment"
     repeat_region     26428..26491
                     /note="MSTR element fragment"
     repeat_region     26465..26518
                     /note="MSTR element fragment"
     repeat_region     27446..27667
                     /note="L1 element fragment"
     repeat_region     28551..29247
                     /partial
     repeat_region     30160..30325
                     /note="Alu repeat: matches 308. .1 of consensus"
     misc_feature     32588..32875
                     /note="2 copies of 83 mer 98 & conserved"
     misc_feature     32588..32875
                     /note="match: 5' EST H68599 clone 239077"
     repeat_region     35531..35802
                     /note="match: 3' EST H68239 clone 289077"
     repeat_region     35531..35802
                     /partial
     repeat_region     35804..35939
                     /note="Alu repeat: matches 1. .260 of consensus"
     repeat_region     38625..38705
                     /note="18 copies of 2 mer 83 & conserved"
     repeat_region     38747..38839
                     /note="MSTR element fragment"
     repeat_region     38949..39032
                     /note="MSTR element fragment"
     repeat_region     38965..39076
                     /note="MSTR element fragment"
     repeat_region     38965..39076
                     /note="MSTR element fragment"
BASE COUNT  14260 a 9135 c 9327 g 12956 t
ORIGIN

```

Query Match 20.8%; Score 185.2; DB 9; Length 45678;
 Best Local Similarity 64.4%; Pred. No. 5.2e-40;
 Matches 345; Conservative 0; Mismatches 178; Indels 13; Gaps 4;

```
/note="LMD repeat: matches 1357. .1972 of consensus"
11443. .11571
repeat_region
/note="FLAM A repeat: matches 2. .129 of consensus"
11572. .11946
repeat_region
/note="LMD repeat: matches 977. .1357 of consensus"
11947. .12245
repeat_region
/note="AluY repeat: matches 1. .301 of consensus"
12246. .11098
repeat_region
/note="LMD repeat: matches 275. .977 of consensus"
12453. .14334
repeat_region
/note="LMD repeat: matches -17. .113 of consensus"
13523. .13820
repeat_region
/note="MTRD repeat: matches 177. .505 of consensus"
13823. .13932
repeat_region
/note="55 copies 2 mer ct 77% conserved"
13936. .13977
repeat_region
/note="MTRD repeat: matches 201. .241 of consensus"
13978. .14240
repeat_region
/note="AluSg1 repeat: matches 1. .270 of consensus"
14245. .14334
repeat_region
/note="AluSg1 repeat: matches 1. .290 of consensus"
14535. .14921
repeat_region
/note="MTRD repeat: matches 2. .201 of consensus"
14724. .15604
repeat_region
/note="LMD repeat: matches 4646. .5565 of consensus"
15605. .15909
repeat_region
/note="AluSx repeat: matches 24. .312 of consensus"
16069. .16363
repeat_region
/note="AluSx repeat: matches 1. .291 of consensus"
16482. .16641
repeat_region
/note="AluDb repeat: matches 136. .295 of consensus"
16711. .18293
repeat_region
/note="AluSg1 repeat: matches 1. .284 of consensus"
17164. .17280
repeat_region
/note="LIP repeat: matches 4499. .4619 of consensus"
17326. .17377
repeat_region
/note="26 copies 2 mer at 79% conserved"
17704. .17863
repeat_region
/note="AluSc repeat: matches 118. .295 of consensus"
17864. .18175
repeat_region
/note="AluYb8 repeat: matches 1. .318 of consensus"
18176. .18298
repeat_region
/note="AluSc repeat: matches 1. .118 of consensus"
18308. .18379
repeat_region
/note="113 copies 2 mer tg 100% conserved"
18876. .19171
repeat_region
/note="AluSg repeat: matches 1. .295 of consensus"
19183. .19232
repeat_region
/note="25 copies 2 mer ac 100% conserved"
19246. .19533
repeat_region
/note="AluY repeat: matches 5. .292 of consensus"
19539. .19837
repeat_region
/note="MER748 repeat: matches 61. .368 of consensus"
2016. .20274
repeat_region
/note="MER748 repeat: matches 295. .487 of consensus"
20394. .20592
repeat_region
/note="AluSg repeat: matches 1. .302 of consensus"
20701. .20809
repeat_region
/note="LIR1 repeat: matches 682. .785 of consensus"
20877. .21061
repeat_region
/note="LIR1 repeat: matches 1. .195 of consensus"
21186. .21636
repeat_region
/note="LIR29 repeat: matches 3. .498 of consensus"
21609. .21737
repeat_region
/note="MER34 repeat: matches 409. .543 of consensus"
1771. .12076
repeat_region
/note="LIR repeat: matches 2. .308 of consensus"
22878. .22977
repeat_region
/note="MIR repeat: matches 92. .191 of consensus"
23452. .23582
repeat_region
/note="FLAM C repeat: matches 1. .131 of consensus"
23583. .23624
repeat_region
/note="21 copies 2 mer ta 76% conserved"
```

```
23735. .24041
repeat_region
/note="AluSg1 repeat: matches 1. .308 of consensus"
24045. .24072
repeat_region
/note="14 copies 2 mer tt 89% conserved"
25274. .25582
repeat_region
/note="AluDb repeat: matches 1. .289 of consensus"
25583. .25747
misc_feature
/note="match: GSS B63622"
25959. .26012
repeat_region
/note="HERVFR1 repeat: matches 23. .79 of consensus"
26144. .26436
repeat_region
/note="AluSg1 repeat: matches 1. .311 of consensus"
27281. .27432
repeat_region
/note="AluYb repeat: matches 137. .289 of consensus"
27436. .27507
repeat_region
/note="LIM4 repeat: matches 5360. .5434 of consensus"
27536. .27600
repeat_region
/note="LIM4 repeat: matches 2220. .2283 of consensus"
27601. .27975
repeat_region
/note="MTRD repeat: matches 3. .365 of consensus"
27976. .28273

Query Match      17.8%   Score 158.6;   DB 9;   Length 59648;
Best Local Similarity 60.1%;   Pred. No. 1,3e-32;
Matches 409;   Conservative 0;   Mismatches 229;   Indels 43;   Gaps 7;

QY 3 cccatccccccactctatccgctctccatttgggtgctgcaagctctgggaaagaa 62
Db 64012 CCCCTTCCACACGCCATCTCGCTCTCATTTTGTCTGCTCGCGAAGCCTGGAGTGA 64071

QY 63 tcccggaacacgaaaatggtggtttgggggaaggaggaaggaaggaagctggag 122
Db 64072 TC---TACGGGAATATGTGGCTTTCGGGAAGGAGGGCTCGGATTCAGGCGACCC 64127

QY 123 gagggtttaaaggagccctatagagagcgcgcggaaacttctaaggtgggaaaa-- 180
Db 64128 CACAGGCNACTGCTCTCAGTGGGAAATAAATGTTAGATATTAAGTCAAGT 64187

QY 181 -----acgaataataaaatctcttgatatacagggtctgaatcctgctggtcagag 232
Db 64188 CCAGGCTCTCAATTTAAAGGTGGCTAGTAGGGCTCTGTCTCAGTGTCTATCAGTC 64247

QY 233 ccccaagcattcgtctctctctctctctctctctctctctctctctctctctctctct 287
Db 64248 CACCAAGATTAAGCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 64307

QY 288 caacagaaaaaaatctctatcagcaaatattcacgagaaacagagagatgag 347
Db 64308 GAACATAGCGCAACATCTCAACGGGAAAATGCCCAACAGAAACAGAGAGGGGAG 64367

QY 348 cag---ctatgcagaatgagagagacgcgcctttggggagaggtgaagccaccag 404
Db 64368 CAGCCCTCAGCGAATGAGAGAAATCCCGCCATTGGGAGGGGGTGAAGGCCAAG 64427

QY 405 cctcgaggaat---cgacggggacaggtgcgcacttgccttaatttcgagggcc 461
Db 64428 CTGGAGGAATATCAGCGGGGGCGAGTTAGCGGCTGTCTCTATTTTCATTTTCATGGGCC 64487

QY 462 atccccatagcgaatcaatgctgggatgggtgagatgagagatgagaaatttc 521
Db 64488 ATACCTAATAGGCATAT-----TGAGCACAATGAAGCGAGAGATGATGAGAAAGGTTT 64541

QY 522 atgggagagatgagagaaatcagaaacacttggagagctgagtgaggaattgtctg 581
Db 64542 GTAGGGCAGATGATGGAATCAGAGAAAGACTAGGGAACAGCAGATGAGGACACTATATG 64601

QY 582 cgtctccttatgggggggctctcctcaatcacatgacatcatgatgaatttgccttatg 641
Db 64602 CCTTCCAAA-----CCTCGACCTGACACCATATATCTTTTCTCTCTA 64649

QY 642 ccttgactcctgcattatc 662
Db 64650 CTTTGAATCTTAAAGTTTTC 64670
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using av model

Run on: March 11, 2002, 16:09:50 ; Search time 1609.39 seconds
11351937 seqs, 537289281 residues
5949.141 Million cell updates/sec

Title: US-09-327-750d-29

Perfect score: 891
Sequence: 1 accaccatccccctctat.....ataaagcaatttaaaagc 891

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

- 1: em_estfun:**
- 2: em_esthum:**
- 3: em_estlin:**
- 4: em_estom:**
- 5: em_estpl:**
- 6: em_estba:**
- 7: em_estro:**
- 8: em_estov:**
- 9: em_hcc:**
- 10: gb_estl:**
- 11: gb_estc:**
- 12: gb_estd:**
- 13: gb_gss:**
- 14: em_gss_fun:**
- 15: em_gss_hum:**
- 16: em_gss_inv:**
- 17: em_gss_pln:**
- 18: em_gss_pro:**
- 19: em_gss_rpd:**
- 20: em_gss_vit:**
- 21: em_gss_other:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	607.8	68.2	612	10	AW512400
C 2	597.0	67.7	615	10	AI193112
C 3	597.4	67.9	626	10	AI193112
C 4	586.4	66.9	626	10	AI193112
C 5	577.4	64.8	586	11	AF058976
C 6	577.4	64.8	586	10	AF058976
C 7	574	64.4	662	10	AI1929106
C 8	565.2	63.4	603	10	AA514324
C 9	564.4	63.3	568	10	AI131584
C 10	564.4	63.3	574	10	AI1075989
C 11	564.4	63.3	610	10	AV709328
C 12	562.8	63.2	574	10	BE222704

AW512400
xx74b09.x1 NCI-CGAP-Lym12 Homo sapiens cDNA clone IMAGE:2849369 3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN). mRNA sequence.
AW512400
EST:400.1 GI:7150478
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: rstraus@fmail.nih.gov
Lab: Research by NCI-CGAP at the University of Maryland
DNA Sequencing by NCI-CGAP Consortium/LLNL at:
found through the I.M.A.G.E. Consortium/LLNL at:
Image.llnl.gov/image/html/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence, stop: 422.
Location/Qualifiers
1. 612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2849369"
/clone_lib="NCI-CGAP-Lym12"
/tissue="lymphoma, follicular mixed small and large cell"

ALIGNMENTS

RESULT 1
AW512400/c
LOCUS
DEFINITION
xx74b09.x1 NCI-CGAP-Lym12 Homo sapiens cDNA clone IMAGE:2849369 3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN). mRNA sequence.
ACCESSION
AW512400
VERSION
KEYWORDS
EST:400.1 GI:7150478
SOURCE
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: rstraus@fmail.nih.gov
Lab: Research by NCI-CGAP at the University of Maryland
DNA Sequencing by NCI-CGAP Consortium/LLNL at:
found through the I.M.A.G.E. Consortium/LLNL at:
Image.llnl.gov/image/html/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence, stop: 422.
Location/Qualifiers
1. 612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2849369"
/clone_lib="NCI-CGAP-Lym12"
/tissue="lymphoma, follicular mixed small and large cell"

```

Oy 821 aegttctgcagctagttctcccccatttcgaggaataattaaagcaataaagca 880
Db 75 AAGTTCTGTCAGCAGTAGTTCTACCCATTTCGATCGGAAATATTCCTCCCTAATAAGCA 16
Oy 881 atttaaaagc 891
Db 15 ATTTAAAGC 5

RESULT 3
A1928835/c 626 bp mRNA EST 23-AUG-1999
LOCUS eu55b10.x1 schneider fetal brain 00004 Homo sapiens cDNA clone
DEFINITION IMAGE:2519611.3 similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
ACCESSION A1928835 IN HGR74 (HUMAN); mRNA sequence.
VERSION A1928835.1 GI:5664828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier, L., Allen, M., Bowles, L., Dubouche, T., Gelsel, G., Jost, S.,
Kritman, D., Kucaba, T., Le, N., Lennon, G., Marra, M., Martin,
M., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
Wu, J., Wyllie, J., Yeh, W., Zeng, L., Wilson, R., and Wilson, R.
TITLE Washington University School of Medicine
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevan@wustl.edu
This clone is available royalty-free through LNC. contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Gibco
High quality sequence stop: 475.
FEATURES
Location/Qualifiers
1..626
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2519611"
/clone_lib="Schneider fetal brain 00004"
/sex="Male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGCTCAAGATCTTAAATTAATTCCTCCCTCCCTCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGAGCTCAAGATCTTAAATTAATTCCTCCCTCCCTCC-3'
The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."
BASE COUNT 175 a 144 c 122 g 184 t
ORIGIN
Query Match 67.0%; Score 597.2; DB 10; Length 626;
Best Local Similarity 98.5%; Pred. No. 5.3e-114;
Matches 602; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Oy 281 agaaaaaacaggaataaaatctcatcatcattggaataattaccaggaataaagca 340

```

[illegible]

[illegible][illegible]

AI075989.1 GI:3405167
 EST.
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 574)
 NC-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapsb-remail@nci.nih.gov
 DNA LIBRARY PREPARATION: M. Bento Soares, Ph.D., M. Fatima Bonaldo


```

OY 806 ctttgcaactcagtgtaagttcttcgcagcagtagttccaccatttgcatggaaatttt 865
DB 86 ctttgcaactcagtgtaagttcttcgcagcagtagttccaccatttgcatggaaatttt 27

OY 866 aagccataaagcaatttttaaaagc 891
DB 26 AAAGCTAATAAGCAATTTTAAAGAC 1

RESULT 15
BF732486/c 566 bp mRNA EST 30-MAR-2001
LOCUS naelid11.x1 NCI-CGAP-Ov18 Homo sapiens cDNA clone IMAGE:343488 3'
DEFINITION similar to SW:HG74, HUMAN O00994 OVARIAN GRANULOSA CELL 13.0 KD
PROTEIN HG74.1, mRNA sequence.
ACCESSION BF732486
KEYWORDS
SOURCE Human
ORGANISM Homo
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 566)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Samuels, Ph.D., Ph.D. cDNA library: Reproductive Medicine
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 453.
FEATURES
source
1..566
/organism="Homo sapiens"
/db_xref="GeneID:343488"
/clone_lib="NCI-CGAP-Ov18"
/tissue_type="fibrotheca"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Ovary; Vector: pTZ19-Pac (Pharmacia) with a
modified polylinker; Site.1: Not 1; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCACTCTGAAGTGGAGCGCCGCCGACATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pTZ19 vector. Library
constructed by Banco Soares and M.Fatima Bonaldo."
BASE COUNT 162 a 140 c 110 g 153 t
ORIGIN

```

Search completed: March 11, 2002, 16:09:59
Job time: 10894 sec

```

Query Match 62.9% Score 560.8; DB 11; Length 566;
Best Local Similarity 99.5%; Pred. No. 1.9e-106;
Matches 562; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 327 caggaacacgaagatgagcagcctatcgacatggagagagaccgccctttgga 386
DB 566 CAGGAACACGAAGATGAGCAGCCTATCGACATGGAGAGAGACCCCTTTGGG 507

OY 387 ggaagtgaaggccagcctcgaggaatcgagggagagctgcgcacttgcct 446
DB 506 GAAGGTGAAGGCCACCAAGCTCGACGAATCGAGGGAGAGCTGCCGACTTGCCT 447

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sv model

Run on: March 11, 2002, 16:33:56 ; Search time 61.41 Seconds
(without alignments)
3283:970 Million cell updates/sec

Title: US-09-327-750D-29
Perfect score: 891
Sequence: 1 accctaccctccactctat.....ataaagcaatttaaaagc 891

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	62.4	7.0	7218	1	US-08-232-463-14
C 2	47.8	5.4	7218	1	US-08-232-463-14
C 3	44.4	5.0	40328	3	US-08-742-185-102
C 4	37.2	4.2	1871	1	US-08-416-478A-1
C 5	37.2	4.2	1871	2	US-08-474-988B-1
C 6	37.2	4.2	1871	2	US-08-394-442B-1
C 7	37.2	4.2	289	4	US-08-007-005-17
C 8	37.2	4.2	289	4	US-08-454-708-12
C 9	36.6	4.1	581	3	US-08-334-306-22
C 10	36.6	4.1	581	3	US-08-334-306-22
C 11	36.6	4.1	581	3	US-08-334-306-22
C 12	35.8	4.0	278	2	US-08-332-765A-42
C 13	34	3.8	2277	1	US-08-576-967-2
C 14	34	3.8	2277	1	US-08-576-967-2
C 15	34	3.8	2277	2	US-09-098-487-2
C 16	33	3.7	53526	3	US-08-558-136-2
C 17	33	3.7	53577	3	US-08-558-136-2
C 18	32.2	3.6	4317	4	US-08-488-706-2
C 19	32.2	3.6	4317	4	US-08-488-706-2
C 20	32.2	3.6	4317	4	US-08-488-706-2
C 21	32.2	3.6	5681	1	US-09-007-006-17
C 22	32	3.6	289	4	US-09-244-796-17
C 23	32	3.6	289	4	US-09-244-796-17
C 24	32	3.6	623	4	US-09-385-982-167
C 25	32	3.6	246240	2	US-08-724-394A-20
C 26	32	3.6	246240	2	US-08-724-394A-21
C 27	32	3.6	246240	2	US-08-724-394A-22

28	31.8	3.6	772	2	US-08-989-337-2
29	31.8	3.6	13146	2	US-08-724-354D-3
30	31.8	3.6	13146	3	US-09-270-984A-3
C 31	31.6	3.5	24979	2	US-08-147-777-3
C 32	31.6	3.5	24979	3	US-08-452-872-3
C 33	31.6	3.5	24979	5	PCT-US93-03985-3
C 34	31.4	3.5	990	3	US-08-921-209-3
C 35	31.4	3.5	990	3	US-08-411-763-3
C 36	31.4	3.5	1000	2	US-08-927-615-1
C 37	31.2	3.5	3583	4	US-08-211-704A-1
C 38	31.2	3.5	3583	3	US-08-048-889-6
C 39	31	3.5	3583	3	US-08-048-889-6
C 40	31	3.5	13146	2	US-08-724-354D-3
C 41	31	3.5	13146	3	US-09-270-984A-3
C 42	31	3.5	87350	3	US-08-781-891-79
C 43	30.8	3.5	3356	4	US-08-974-549A-640
C 44	30.8	3.5	6678	3	US-08-816-617A-1
C 45	30.6	3.4	32207	2	US-08-770-379-20

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOULFOX VIRUS
; NUMBER OF SEQUENCES: 32
; NUMBER OF SEQ. POSITIONS: 52
; COMPLETION DATE: 1998-08-26
; ADDRESSER: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA: Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91.114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; FIRM: BENT, Stephen A., 9, 768
; REFERENCE/DOCUMENT NUMBER: 30,772/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 895149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; ORGANISM: FOULFOX VIRUS
US-08-232-463-14

Sequence 2, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 6, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 79, Appl
Sequence 640, Appl
Sequence 1, Appl
Sequence 20, Appl

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 861-8540
 TELEFAX: (617) 861-8540
 INFORMATION FOR SEQ ID NO: 102:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 40328 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-742-185-102

Query Match 5.0%; Score 44.4; DB 3; Length 40328;
 Best Local Similarity 51.0%; Pred. No. 0.012;
 Matches 105; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
 Oy 684 cegctgtttctttcttctgattcttaataatgctttactgacgtctgtcgtgga 743
 Db 20178 CCCCTCCCTTTTTTTTTCCTTCCTGCTGAGATCTTCCTGCATCTTTTATTGGA 20237
 Oy 744 accctatgtattcttcagtgctcaagtggtgctgtgctcagctctctattgaagt 803
 Db 20238 GCTATGTGTGTCCTGACGAGATGGTTCCTGATACAGCAGCTGATGGGTCTT 20297
 Oy 804 gctttgactcagctgaagttctctcagcagctagttctaccctattgctggaagt 863
 Db 20298 GACTCTTTATCCAGTTGGCCAGTCTGTGCTTTTAAATGGAGCATTTAGTCCATTCACN 20357
 Oy 864 ttaagcaataaagcaatttaaaaa 889
 Db 20358 TTAAGTTAATATCTTATGATGAA 20383

RESULT 4

US-08-416-478A-1/c
 Sequence 1, Application US/08416478A
 Patent No. 5755757
 GENERAL INFORMATION:
 APPLICANT: Hercend, Thierry
 APPLICANT: Hercend, Thierry
 TITLE OF INVENTION: New Proteins Produced by Human
 TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
 TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROMDY AND NEIMARK, P.L.L.C.
 STREET: 419 Seventh Street, N.W., Suite 400
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,478A
 FILING DATE: 04-APR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/854,644
 FILING DATE: 08-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: YUN, Allen C.
 REGISTRATION NUMBER: 37,971
 REFERENCE/DOCKET NUMBER: HERCEND-1A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1871 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 231..1724

LENGTH: 1871 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 231..1724
 US-08-416-478A-1
 Query Match 4.2%; Score 37.2; DB 1; Length 1871;
 Best Local Similarity 57.9%; Pred. No. 0.29;
 Matches 66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 Oy 14 ctcctatccggtctctccattttgtgctgctgcaagctctggaagaatccccgggaac 73
 Db 239 CCCCACATCTCTCTATGCTCTGGCAGCCCTCGAGCCCTCCCAAAAGAGGTCAGAAAG 180
 Oy 74 gaaaatgtgctttgaggaagaagagctaaagggagaaagctgagggaggg 127
 Db 179 GAAAGGGGCGAGTGGTGGGGCTGGGGGTAAAGGAGAGCTCTGCGAGAGGG 126
 RESULT 5
 US-08-474-988B-1/c
 Sequence 1, Application US/08474988B
 Patent No. 5874250
 GENERAL INFORMATION:
 APPLICANT: Hercend, Thierry
 APPLICANT: Hercend, Thierry
 TITLE OF INVENTION: New Proteins Produced by Human
 TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
 TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROMDY AND NEIMARK, P.L.L.C.
 STREET: 419 Seventh Street, N.W., Suite 400
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,988B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/416,478
 FILING DATE: 04-APR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/854,644
 FILING DATE: 08-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: YUN, Allen C.
 REGISTRATION NUMBER: 37,971
 REFERENCE/DOCKET NUMBER: HERCEND-1B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1871 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 231..1724

```
Query Match      4.19; Score 36.6; DB 2; Length 581;
Best Local Similarity 57.48; Pred. No. 0.23;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
Query Match      4.18; Score 36.6; DB 2; Length 581;
Best Local Similarity 57.48; Pred. No. 0.23;
Matches 66; Conservative 0; Mismatches 49; Indels
```

[illegible]

RESULT 10
 US-08-814-306-22/C
 : Sequence 22, Application US/08834306
 : Patent No. 6054135
 : GENERAL INFORMATION:
 : APPLICANT: Reed, Steven G.
 : APPLICANT: Skelky, Tasir A.W.
 : APPLICANT: Lodes, Michael J.
 : APPLICANT: Houghton, Raymond L.
 : TITLE OF INVENTION: COMPOUNDS AND
 : NUMBER OF SEQUENCES: 65
 : CORRESPONDENCE ADDRESS:
 : DR. STEVEN G. REED AND BERRY LLP
 : STREET: 7700 Columbia Center, 7
 : CITY: Seattle
 : STATE: Washington
 : COUNTRY: USA
 : ZIP: 98104-0992
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release 11.0
 : CURRENT APPLICATION DATA: 08/834,3
 : FILING DATE: 15-SEP-1997
 : CLASSIFICATION: 424
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Makl, David J.
 : REGISTRATION NUMBER: 31,392
 : REFERENCE/POCKET NUMBER: 210121
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (206) 522-4900
 : TELEFAX: (206) 682-6031
 : INFORMATION FOR SEQ ID NO: 22:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 591 base pairs
 : TYPE: nucleic acid
 : STRATEGY: single
 : TOPOLOGY: linear
 : US-08-814-306-22

Query Match 4.18; Score 36.6; DB 3; Length 581;
Best Local Similarity 57.48; Pred. NO. 0.23;

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,967
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-676-967-2

Query Match 3.8%; Score 34; DB 1; Length 2277;
Best Local Similarity 29.0%; Pred. No. 2.8;
Matches 54; Conservative 38; Mismatches 94; Indels 0; Gaps 0;

QY 572 gaattgtctgcgtatccttctatggggagctctcttaaccatgaccatcatgatgaatt 631
DB 2224 RNATYTTGTTTATGTCACNNAATGTTTAAACNGKNGTTCNGTTCNCCT 2165
QY 632 ttccctatgcctgcctgcctgcattatcatcatgatgaattatcatgcctgcctgtt 691
DB 2164 TNGCYTTTTCNCKNNACATGTCNSNSNARYTGYTGYTTCYTTCCAYTGRT 2105
QY 692 tctctttctgcattctctatcatgccttcatgcctgttgcctggaacctatg 751
DB 2104 TDATYNGGCTTNGGCTTTCNGRTGACNGGCTTTCNCCTTTCNCKNNAC 2045
QY 752 tttatt 757
DB 2044 KDATYTT 2039

RESULT 14
US-08-676-974-2/c
Sequence 2, Application US/08676974
Patent No. 5770422
GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,974
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-098-487-2

TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-676-974-2

Query Match 3.8%; Score 34; DB 1; Length 2277;
Best Local Similarity 29.0%; Pred. No. 2.8;
Matches 54; Conservative 38; Mismatches 94; Indels 0; Gaps 0;

QY 572 gaattgtctgcgtatccttctatggggagctctcttaaccatgaccatcatgatgaatt 631
DB 2224 RNATYTTGTTTATGTCACNNAATGTTTAAACNGKNGTTCNGTTCNCCT 2165
QY 632 ttccctatgcctgcctgcctgcattatcatcatgatgaattatcatgcctgcctgtt 691
DB 2164 TNGCYTTTTCNCKNNACATGTCNSNSNARYTGYTGYTTCYTTCCAYTGRT 2105
QY 692 tctctttctgcattctctatcatgccttcatgcctgttgcctggaacctatg 751
DB 2104 TDATYNGGCTTNGGCTTTCNGRTGACNGGCTTTCNCCTTTCNCKNNAC 2045
QY 752 tttatt 757
DB 2044 KDATYTT 2039

RESULT 15
US-09-098-487-2/c
Sequence 2, Application US/09098487
Patent No. 5917025
GENERAL INFORMATION:
APPLICANT: COLLINS, Kathleen
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,487
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-098-487-2

gb_ba:BAFTMX	+	93.50	137.19	22.97	1315	X74118 B.bronchiseptica f
gb_pat:A73843	+	93.50	137.19	22.97	1315	X74118 B.bronchiseptica f
gb_pat:A73843	+	93.50	137.19	22.97	1315	X74118 B.bronchiseptica f
gb_pat:AF152500	-	93.00	132.03	44.54	2382	AF152500 Homo sapiens ptc
gb_pat:AF152500	-	93.00	132.03	44.54	2382	AF152500 Homo sapiens ptc
gb_pat:AK024641	-	93.00	130.73	52.61	2837	AK024641 Homo sapiens cdc
gb_pat:AF217750	-	93.00	129.08	65.08	3547	AF217750 Homo sapiens cdc
gb_pat:AF217750	-	93.00	129.08	65.08	3547	AF217750 Homo sapiens cdc

seq.documentation_block: 11-APR-1999 :
LOCUS AF097438 835 bp mRNA
MUS musculus brain expressed X-linked protein 1 (Bxl)¹, mRNA,
complete cds.
ACCESSION AF097438
VERSION AF097438.1 GI:4580589
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1. Brown, A.L. and Kay, J.C.F.
Title: Bxl, a member of a novel gene family on the mouse X chromosome
Journal Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
ERRATUM Erratum: [[published erratum appears in Hum Mol Genet 1999
May; 8(5):943]]
REMARK 2 (bases 1 to 835)
REFERENCE

```

JOURNAL
Submitted (08-06-2023) Cancer Unit, Queensland Health Institute of Medical Research, Brisbane, Qld 4039, Australia
FEATURES
source
1..835
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"
/map="near Pip"
/tissue_type="blastocyst"
1..835
gene
1..835
cds
206..591
/codon_start=1
/gene="Rex1"
/product="brain expressed X-linked protein 1"
/protein_id="AA024429.1"
/db_xref="GI:4580590"
/translocation="MESKDQGVKNLNMENDHOKKEKPEKPDTIRREPAVALISBAG
KNCAPGRGRRRVROPTINRYNDRDGLVLYSAANCVYALPQARG
QLSHLSAVNSTDPFHEMDHFCLAP"
BASE COUNT      238 a   104 c   246 g   167 t
ORIGIN

alignment_scores:
Quality: 688.00           Length: 128
Ratio: 5.375             Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000..100

alignment_block:
US-09-327-7500-30 x AF097438 ..
Align seq 1/1 to: AF097438 from: 1 to: 835

1 MetGlusterLyaspGInGlyValLyAsanLeuanMetGluasnasPHI 17
|||||
17 ATFGAGTCTCAACGTATCAGGGCGTGAAAATCTCACATCATGAGAATGCCA 254
|||||
205 scILyLyvscGltGlycGluGlypsrGclnagshprltsatgacg 34
|||||
215 TCAGNAANAAGGACGAGGAAGGAGGAAGAACCCACACATCATCAGAGGG 304
|||||
34 luProAlavAlaLeuIfesTerGluAGlGlyVsancVyAlaPePrArg 50

```

33 rgluproalvalaleuileSerGlualeGlyAsnCysAlaPro 49
100
264 GGGAGCC...TTGGCCCTCCCTTGGATGCTGGTGAATACTGTGTGCT 310
50 ArgGlyValArgArgPheArgValArgGlnProIleAlaHisTyr 66
100
311 AGAGAAATCGTAGCGGTTCCGGCTTAGCGAGCCATCTCGCAGTATAG 360
66 gtrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGlu 83
100
361 ATGGGATATGATGATAGCTTAGCTTGGAGACCCACAGCAGGATGAGAGAG 410
83 luAnValGlnArgPheGlyValArgGlnLeuMetGluLysLeu 99
100
411 ACAATATGAAGGATGGGGAGGGGTGAGACACCTGATGGAAGCTG 460
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
100
461 AGGAAAGACAGTAGTGTAGTCTAGTCTGGGGCAGTACAGCAGTACCC 510
116 ohIshIsAspHisAspGluPheCysLeuMetPro 128
100
511 TCNCCATGACCTCATGATGATGATTTGCTTTGCCCC 547
seq_name: /SDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAH75810
seq_documentation_block:
ID_AA75810 standard; cDNA; 767 BP.
XX AAH75810;
XX
XX 17-OCT-2001 (first entry)
XX Human X chromosome linked gene expression protein 14 coding sequence.
XX Human; X chromosome linked gene expression protein 14; cancer;
XX HIV infection; cytostatic; anti-HIV; chromosome X; ss.
XX Homo sapiens.
XX CN1296969-A.
XX
XX 30-MAY-2001.
XX
XX 23-NOV-1999; 99CN-0124078.
XX
XX 23-NOV-1999; 99CN-0124078.
XX (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
XX
XX Mao Y, Xie Y;
XX WPI: 2001-48397/53.
XX P-PSDB: AAG66407.
XX Polypeptide-human X chromosome linked gene expression protein 14 and
XX polynucleotide for coding said polypeptide -
XX
XX Claim 6; Page 24 (Disclosure); 31pp; Chinese.
XX The present sequence is the coding sequence for human X chromosome linked
XX gene expression protein 14. The protein and coding sequence are useful
XX for treating diseases e.g. cancer and HIV infection.
XX
XX Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;
XX
alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:
US-09-327-750D-30 x AAH75810 ..
Align seg 1/1 to: AAH75810 from: 1 to: 767
1 MetGlySerLysAspGln...GlyValLysAsnLeuAsnMetGluAsn 16
100
101 ATGCAGTCCAAAGAGAACGAGCGTTAAACACATCTCATCTGCGGAAATGT 150
16 PHisGlnArgValGluGluArgGlnGluArgPheGlnAspThrIleAsn 33
100
151 CAACCAAGAAATGATGAAAGATGAAAGAGCAAGTTGCTAATAAG 200
33 rgluproalvalaleuileSerGlualeGlyAsnCysAlaPro 49
100
201 GGGAGCC...TTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCT 247
50 ArgGlyValArgArgPheArgValArgGlnProIleAlaHisTyr 66
100
248 AGAGAAACCGTAGCGGTTCCGGCTTAGCGAGCCATCTCGCAGTATAG 297
66 gtrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGlu 83
100
298 ATGGACATATGATGATGATGATTTGCTTTGCCCC 347
83 luAnValGlnArgPheGlyValArgGlnLeuMetGluLysLeu 99
100
348 AGAATATGAAAGATGGGGAGGGGTGAGACACCTGATGGAAGCTG 397
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
100
398 AGGAAAGACAGTAGTGTAGTCTAGTCTGGGGCAGTACAGCAGTACCC 447
116 ohIshIsAspHisAspGluPheCysLeuMetPro 128
100
448 TCACCATGACCATCAGTATGATTTGCTTTGCCCC 484
seq_name: /SDS2/gcdata/geneseq/geneseq/NA2001.DAT:AA158581
seq_documentation_block:
ID_AA158581 standard; cDNA; 862 BP.
XX AA158581;
XX
XX 22-OCT-2001 (first entry)
XX Human polynucleotide seq ID NO 784.
XX
XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX ankyrotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX Homo sapiens.
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 23-APR-2000; 2000US-0524317.
XX
XX 19-JUL-2000; 2000US-0520412.
XX
XX 03-AUG-2000; 2000US-0620312.
XX
XX 14-SEP-2000; 2000US-0653450.
XX
XX 19-OCT-2000; 2000US-0652191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX

claim 1: SEQ ID 14964: 71bp + CD-ROM: English.

The present sequence is one of a large number of 5' ESTs derived from

total human RNAs or polyA+ RNAs derived from 30 different tissues. EST

of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs, and in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are commonly used in diagnostic, forensic, gene therapy and chromosome mapping studies. They are also used to identify regulatory sequences and to design expression vectors.

```

alignment_scores:      429.50      Length: 125
                       Quality: 3.940      Gaps: 1
                       Percent Similarity: 87.200      Percent Identity: 64.000

alignment_block:
US-09-327-750D-30 x AAC10889      ..      alignment seg 1/1 to: AAC10889      from: 1 to: 592

4      LysaSGcngLyvAlLysaSLenUanNetcJLaaNAspHisGlnLyLyL 20
249      AGAGACAGGCGTTAAACAATCTCATCGTGGMAAAATGTCACACGAGAAA 298

```

[illegible]

299 TGATGAAAAGATGAAAGGAGCAAGTGWCTAATAAAGGGGAGCCC...T 345

SECRET

305 ACCCTTCCCCCTTACCGAGCCCATCTCCAGTATGATGGGATATGAT 445

70 tGlnArgValGlyGluProGlnGlyArgMetArgGluAsnValGlnA 87

446 GCATAGGCTTGAGAACCCACAGGCAAGGATGANAGAGAGAAATATGGAAA 495

87 rgpheGlyGlyAspValArgGlnLeuMetGluLysLeuArgGluArgGln 103

496 GGATTGGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGGGAAAAGCAG 543

104 LeuserHlsserLeuAArgAlaValSerThrAspProHishHisAspHt 120
|||||

```
1 MetGluSerIysAsp...GlnGlyValIysAsnLeuAnMetGluAsnAs 16
24358 ATGGATGCTCAAGAGAAAGAGCAGCTAAACAGCTCTCAGCATGGAAATGCG 24309
16 pHisGlnIysGlySgIuGluLysProGlnAspThrIleArgA 33
24308 CACCAGAAATATGAGAAAGAG...CAAGTTGCTTAATAAAG 24268
33 TGGTUPGAlaValAlaLeuIleSerGluAlaGlyIysAsnCysAlaPro 49
24267 GGGAGGCC...TTGGCCCTCCCTTTGGATGCTGTGTAATCTGTGCT 24221
50 ArgGlyIysArgArgArgPheArgValArgGlnProIleAlaHisIyAr 66
24220 AGAGAAATCTAGGGGGTTCCCGCTTAGGCAGCCATCCTGCGATATAG 24171
66 gTTPAspLeuMetGlnArgValGlyGluProGlnIysArgMetArgGluG 83
24170 ATGGATATGCTATGCTAGCTTGGAGAACCCAGGAGGATGAGAGAG 24121
83 IuAsnValGlnArgPheGlyIysAspValArgGlnLeuMetGluLysLeu 99
24120 AGAATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAAGCTG 24071
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
24070 AGGAAAGCACTTGAAGTATAGTCTCGGGCAGCTGAGCAGCTGACCC 24021
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
24020 TCACCTGACCATCATGATGAGTTTGGCTTATGCC 23984
seq_name: gb_pr:AF237783
seq documentation_block: 791 bp mRNA PRI 02-SEP-2000
LOCUS AF237783
DEFINITION Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete cds.
ACCESSION AF237783
VERSION AF237783.1 GI:9963898
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 791)
AUTHORS Yang, Q.S., Ying, K., Xie, Y., and Mao, Y.M.
TITLE A Novel Human X-linked Brain Expressed Protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 791)
AUTHORS Mao, Y.M., Xie, Y., Yang, Q.S., Wu, H., Lin, S. and Ying, K.
TITLE Direct Submision
JOURNAL Submitted (21-FEB-2000) Genetic Research Institute, Fudan University, 220 Handan Rd., Shanghai 200433, P.R.China
FEATURES
source Location/Qualifiers
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 791)
/db_xref="taxon:9606"
/chromosome="X"
/map="between DXS990 and DX1059 by RH (Stanford G3)"
1..791
/gene="BEX1"
172..549
/gene="BEX1"
/codon_start=1
/product="brain-expressed protein BEX1"
/protein_id="P66982.1"
/db_xref="GI:9963898.1"
/translation="MESKERRALNNLVNNOENDEKDEKQVANKGEPLALPDAGSYC
VPGNRRFRVPTIYVPHDMLRGLERQNRNENRIGEGVQLMKLREQLS
HSLRVSSTDPDHHDFCLMP"
BASE COUNT 214 a 172 c 218 g 187 t
ORIGIN
```

```
alignment_scores:
Quality: 450.50 Length: 129
Ratio: 3.952 Gaps: 3
Percent Similarity: 88.372 Percent Identity: 68.992
alignment_block:
US-09-327-750D-30 x AF237783
Align seg 1/1 to: AF237783 from: 1 to: 791
1 MetGluSerIysAspGln...GlyValIysAsnLeuAnMetGluAsnAs 16
172 ATGGATGCTCAAGAGAAATAGCAGTAACAGCTCTCAGCATGGAAATGCG 221
16 pHisGlnIysGlySgIuGluLysProGlnAspThrIleArgA 33
222 CACCAGAAATATGAGAAAGAG...CAAGTTGCTTAATAAAG 262
33 TGGTUPGAlaValAlaLeuIleSerGluAlaGlyIysAsnCysAlaPro 49
263 GGGAGGCC...TTGGCCCTCCCTTTGGATGCTGTGTAATCTGTGCT 309
50 ArgGlyIysArgArgArgPheArgValArgGlnProIleAlaHisIyAr 66
310 AGAGAAATCTAGGGGGTTCCCGCTTAGGCAGCCATCCTGCGATATAG 359
66 gTTPAspLeuMetGlnArgValGlyGluProGlnIysArgMetArgGluG 83
360 ATGGATATGCTATGCTAGCTTGGAGAACCCAGGAGGATGAGAGAG 409
83 IuAsnValGlnArgPheGlyIysAspValArgGlnLeuMetGluLysLeu 99
410 AGAATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAAGCTG 459
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
460 AGGAAAGCACTTGAAGTATAGTCTCGGGCAGCTGAGCAGCTGACCC 509
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
510 TCACCTGACCATCATGATGAGTTTGGCTTATGCC 546
seq_name: gb_pr:AF251053
seq documentation_block: 744 bp mRNA PRI 15-APR-2001
LOCUS AF251053
DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
ACCESSION AF251053
VERSION AF251053.1 GI:13625167
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 744)
AUTHORS Mao, Y., Xie, Y., Zhou, Z., Zhao, W., Zhao, S., Wang, W., Huang, Y.,
Wang, S., Tang, R., Chen, X. and Wu, C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life
Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.
China
FEATURES
source Location/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
101..487
/codon_start=1
/protein_id="P66982.1"
/protein_id="P66982.1"
/db_xref="GI:13625168"
/translation="MESKERRALNNLVNNOENDEKDEKQVANKGEPLALPDAGSYC
VPGNRRFRVPTIYVPHDMLRGLERQNRNENRIGEGVQLMKLREQLS
HSLRVSSTDPDHHDFCLMP"
BASE COUNT 214 a 172 c 218 g 187 t
ORIGIN
```

CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH13628 and
CC AAH92446 to AAH13632 represent human cDNA sequences; AAH92446 to
CC AAH13632 represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:
Quality: 170.00 Length: 130
Ratio: 1.954 Gaps: 6
Percent Similarity: 66.923 Percent Identity: 34.615

alignment_block:
US-09-327-750D-30 x AAH03517

Align seg 1/1 to: AAH03517 from: 1 to: 865

1 MetCluSerlyAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||.....
191 ATGGAGTCCAAAGAGACTAGTGGCGCAACATCTCAGCGGGGAAATGCG 240
16 PhisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||.....
241 CCACACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
33 TGGGluProAlaValAlaLeuLysSerGluAlaGlyLysAsnCysAlaPro 49
|||||.....
285 AAGAGAAATCCCGCCATTTG.....GGAGCGGCTGAAGCGCAGAGCCT 328
50 ArgGlyGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||.....
329 GGAGGAATATACGGCGCGCGCGAGTATAGCGACTTGTCCCTAATTTTCG 378
66 gtrPaspLeuMetGlnArg...ValGlyGluProGlnGlyArgMetArgG 82
|||||.....
379 ATGGCCCACTACTATAGCATATTCGACCAATGAGCGAGA..... 421
82 LuGluAsnValGlnArgPheGlyAspValArgGlnLeuMetGluLys 98
|||||.....
422 ..GATGATGTAGAAAGTTTGTAGGCGCAGATGATGGAATCAAGAGAAAG 469
99 LeuArgGluArgGlnSerHisSerLeuArgAlaValSerThrAspP 115
|||||.....
470 ACTAGGGAGACGAGTAGAGCGACTATATCGCTTCCCAACTCTCGACCC 519
115 oProHisHisAspHisAspGluPheCysGluPheCysLeuMetPro 128
|||||.....
520 T.....GACACACCTATGACTTTTGGCTCATCT 550

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750
seq_documentation_block:
ID AAH13750 standard; cDNA; 1229 BP.
AC AAH13750;
XX
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:10656.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX

PN EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Nakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX Claim 8: SEQ ID 10656; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises oligo- or primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC particularly full-length cDNA synthesis. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:
Quality: 170.00 Length: 130
Ratio: 1.954 Gaps: 6
Percent Similarity: 66.923 Percent Identity: 34.615

alignment_block:
US-09-327-750D-30 x AAH13750

Align seg 1/1 to: AAH13750 from: 1 to: 1229

1 MetCluSerlyAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||.....
191 ATGGAGTCCAAAGAGACTAGTGGCGCAACATCTCAGCGGGGAAATGCG 240
16 PhisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||.....
241 CCACACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
33 TGGGluProAlaValAlaLeuLysSerGluAlaGlyLysAsnCysAlaPro 49
|||||.....
285 AAGAGAAATCCCGCCATTTG.....GGAGCGGCTGAAGCGCAGAGCCT 328

```

/chromosome="X"
/map="X"
/clone="GHC-870H8"
/clone_lib="SCCV"
475..701
repeat_region /note="L1 element fragment"
repeat_region /note="L1 element fragment"
misc_feature /note="L1 element fragment"
2371..2378
repeat_region /note="match: multiple ESTs"
4189..4246
repeat_region /note="29 copies of 2 mer 91 % conserved"
7496..7874
repeat_region /note="L1 element fragment"
8022..8108
repeat_region /note="L1 element fragment"
8302..8493
repeat_region /note="L1 element fragment"
8585..8671
repeat_region /note="L1 element fragment"
9076..9081
repeat_region /note="L1 element fragment"
9259..9384
repeat_region /note="L1 element fragment"
9624..9731
repeat_region /note="MLTIC element fragment"
9781..9948
repeat_region /note="MLTIC element fragment"
9890..9948
repeat_region /note="MLTIB element fragment"
10181..10392
repeat_region /note="L1 element fragment"
10593..10688
repeat_region /note="Alu repeat: matches 298..108 of consensus"
10688..10781
repeat_region /note="Alu repeat: matches 96..1 of consensus"
11112..11403
repeat_region /note="Alu repeat: matches 1..308 of consensus"
11922..12004
repeat_region /note="L1 element fragment"
12082..12208
repeat_region /note="L1 element fragment"
12423..12707
repeat_region /note="Alu repeat: matches 302..1 of consensus"
12710..12916
repeat_region /note="L1 element fragment"
12979..13081
repeat_region /note="L1 element fragment"
13473..13532
repeat_region /note="MLT2A1 element fragment"
13871..13928
repeat_region /note="L1 element fragment"
14060..14084
repeat_region /note="L1 element fragment"
18778..19056
repeat_region /note="Alu repeat: matches 1..308 of consensus"
22279..22250
repeat_region /note="L1 element fragment"
22449..22895
repeat_region /note="L1 element fragment"
23239..23374
repeat_region /note="L1 element fragment"
23371..23606
repeat_region /note="MLTIB element fragment"
23656..23718
repeat_region /note="MLT2A1 element fragment"
23705..23778
repeat_region /note="MLTIA element fragment"
23805..23960
repeat_region /note="L1 element fragment"

```

```

repeat_region 24011..24541
/note="L1 element fragment"
repeat_region 24638..24770
/note="Alu repeat: matches 150..1 of consensus"
repeat_region 24829..30975
/note="L1 element fragment"
BASE COUNT 9462 a 6147 c 6129 g 5583 t
ORIGIN
alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667
alignment_block:
US-09-327-750D-30 x HSV870H8
Align seg 1/1 to: HSV870H8 from: 1 to: 31321
1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
3508 ATGGAGTCCAAAGAGGACGCGCTTAACATCTCATCTGCGGAATGT 3557
16 pHISGlnLysLysGluGluLysProGlnAspThrIleArg 33
3558 CACCACGAAATCATGAAAGAGGAGGACGAGTTCCTTAATAAG 3607
33 rGluProAlaValAlaLeuLeuSerGluAlaGlyLysAsnCysAlaPro 49
3608 GGGACCC...TTGGCCTACCTTTGAAGTTAGTGAATCTGTGCT 3654
50 ArgGlyGlyValArgArgPheArgValArgGlnProIleAlaHisTyr 66
3655 AGAGGAACCGTAGGGCGGTCCCGTTAGGCAGCCATCTGCAGTAG 3704
66 gTPAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGlu 83
3705 ATGGACATAATGCAAGCTTTGGAGAGCCACGACGAGGATGAGAGG 3754
83 LuanValGlnArgPheGlyAspValArgGlnLeuMetGluLysLeu 99
3755 AGATATGGAAAGCTTGGGAGGAGGTGACACAGCTGATGGAAGCTG 3804
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
3805 AGGAAAGCAGTTGAGTCATAGTCTCGCGGCAGTCAGCAGTCATGCC 3854
116 oHISHisAspHisAspGluPheCysLeuMetPro 128
3855 TCACCATGACCATCAGATGAGTTTGCCTTATGCC 3891
seq_name: gb-pr:AL133348
seq_documentation_block: 40584 bp DNA PRI 06-MAR-2000
LOCUS AL133348 Human DNA sequence from clone RP1-79P11 on chromosome Xq21.32-22.1.
DEFINITION Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed x-linked protein 2), ESTs, STSs, GSSs and a
putative CpG island, complete sequence.
ACCESSION AL133348
VERSION AL133348.8 GI:7076386
KEYWORDS HTG; Bex2; CpG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Wilcox, S.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

```


Quality: 81.00 Length: 121
Ratio: 1.397 Gaps: 4
Percent Similarity: 47.934 Percent Identity: 23.967
alignment_block:
US-09-327-750d-30 x US-08-459-448A-6 ..
Align seg 1/1 to: US-08-459-448A-6 from: 1 to: 3624
14 GluAsnAspHisGlnLysGluGluLysGluLysProGlnAspTh 30
|||||:||||| :|||||:|||||:|||||:|||||:|||||:|||||
1026 GAACACCGCCACCCACCGCCGCCACCAACACCGCATCAACCGCGTGACCC 1075
30 rIleArgArgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnC 47
|||||:||||| :||| :|||||
1076 TCGGCTTCGCCAGCCGCCGCGGTGTACCGCACCGAGAGCTACGCCGCGGTG 1125
47 YeAla.....ProArgGlyArgArgArg 55
:|||||:|||||:|||||:|||||:|||||
1126 CTGCTGTGGGCATCTACCTCGAGCCCATCCACGGCTGCCACCGTGCG 1175
56 PheArgValArgGlnProIleAlaHisTyrArgTrrPAspLeuMetGlnAr 72
:|||||:|||||:||||| :||| :|||
1176 CTTCACTTCACCAACCCGCCAGACATCAGCCA.....CGCG 1213
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
:|||||:|||||:|||||:|||||:|||||
1214 GCACGCCCATCTACGCCACCCCTACGAGGCC..... 1247
89 lYglyAspValArgGlnLeuMetGluLysLeuArgGluArg..... 102
:|||||:|||||:|||||:|||||:|||||
1248CGGCTGCAGCTGAGGACAGCGAGCCGCGTGCCTCC 1286
103GlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
:|||||:|||||:|||||:|||||:|||||
1287 CGAGACACCGAGCGCCCACTACGAGAGCTACAGCCACCGCCTGAGCC 1336
117 sHlAspHisHis 121
||| :|||||
1337 ACNCGGCATCAT 1349

[illegible]

Strd	Orig	zScore	EScore	Len	Documentation
+	683.0	1124.02	1.9e-53	551	! A0522932 Vbl30b28 Beddington m
+	683.0	1122.79	2.2e-53	642	! A0522932 G0104A11-3 N1A Mouse E
+	680.0	1117.07	4.7e-53	707	! BG046920 H3025D01-3 N1A Mouse E
+	670.0	1102.27	3.1e-52	575	! BG277659 ux47c11.x1 Soares_NMMA
+	670.0	1100.27	4.1e-52	756	! BG870503 G0279J432F1 NCI_CGAP_S
+	652.0	1073.78	1.4e-52	499	! BE654459 UT-M-AJ1-aba-f-10-0-UT
+	632.0	1039.69	7.5e-45	577	! BG536974 G0110H01-3 N1A Mouse E
+	591.0	969.71	7.5e-45	785	! BF600762 WY1_000704 Mouse 9-day
-	586.0	963.82	1.6e-44	587	! BG228077 ux47c11.x1 Soares_NMMA
+	542.0	892.26	1.6e-40	583	! BE301079 G0108631F1 NCI_CGAP_M
+	537.0	882.33	5.6e-40	650	! BF608209 WY1_0010188 Mouse 9-day
+	481.5	791.12	6.7e-35	637	! AZ936393 UNO193L05F Mouse 10xb
-	480.5	788.70	9.1e-35	701	! B1183527 NML-P-FN-bf-07-0-UNI
+	476.5	787.27	1.1e-34	370	! D76570 M5U77D10 mouse embryonal
+	471.0	776.11	4.6e-34	480	! AK003429 Mus musculus 18 days e
+	466.5	769.94	1.0e-33	412	! BG659326 DRNAEB11 Rat DRG Libr
+	460.0	756.63	5.6e-33	569	! AV589137 Bos taurus br
-	458.0	750.34	1.2e-32	549	! AW536634 G0106H08-3 N1A Mouse E
+	453.5	746.25	2.1e-32	547	! BG655001 l044f06.y1 HR85 Islet
+	453.5	745.06	2.5e-32	634	! BG742167 G02513894F1 NIH_MGC_1
+	453.5	744.76	2.6e-32	658	! BF971303 G02673150F1 NIH_MGC_9
+	453.5	744.57	2.6e-32	674	! BG744823 G02673146F1 NIH_MGC_8
+	453.5	744.38	2.7e-32	690	! AV020285 AV702885 ADB Homo sapi
+	453.5	744.37	2.7e-32	631	! AV705808 AV705808 ADB Homo sapi
+	453.5	744.32	2.7e-32	695	! A526209 AL526209 LTI_NFL003.NE
+	453.5	744.27	2.7e-32	659	! BG820179 G02782222F1 NCI_CGAP_E
+	453.5	743.36	3.1e-32	783	! BE790774 G01581956F1 NIH_MGC_7
+	453.5	743.25	3.1e-32	794	! A526247 AL526247 LTI_NFL003.NE
+	450.5	739.52	5.0e-32	683	! AV02643 ADB Homo sapi
+	450.0	739.82	4.8e-32	594	! BE314409 G01144007F1 NIH_MGC_9
+	450.0	738.31	5.8e-32	717	! A520931 AL520931 LTI_NFL004.NE
+	450.0	737.71	6.6e-32	712	! BE889313 G01513201F1 NIH_MGC_71
+	450.0	737.32	6.6e-32	870	! BG709503 G02674781F1 NIH_MGC_96
+	450.0	737.08	6.8e-32	835	! BG707398 G02672810F1 NIH_MGC_9
+	450.0	736.07	7.8e-32	946	! BG285825 G02380917F1 NIH_MGC_9
+	450.0	735.29	8.6e-32	1042	! BF038787 G01462119F1 NIH_MGC_3
+	446.5	734.31	9.8e-32	576	! A1929703 au63f04.y1 Schneider
+	446.5	732.35	1.3e-31	734	! BG715659 G02676924F1 NIH_MGC_98
+	446.0	735.57	8.3e-32	445	! BG078231 H3025D01-5 N1A Mouse E
+	444.5	731.73	1.4e-31	537	! BE031524 L30127 MARC 1P1G Sus

ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Minoru S. H. Ko.

BASE COUNT 145 a 182 c 133 g 182 t
ORIGIN
alignment_scores:
Quality: 683.00 Length: 128
Ratio: 5.378 Gaps: 0
Percent Similarity: 99.219 Percent Identity: 99.219

alignment_block:
US-09-327-750D-30 x AW536404/rev

Align seg 1/1 to reverse of: AW536404 from: 1 to: 642

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17
|||||
606 ATGGAGTCCAAAGATCAAGGCGTGAAAAATCTCAACATGGAGATGACCA 557
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArg 34
|||||
556 TCAGAAAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 507
34 lUpProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
506 AGCCACTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 457
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
456 GGAGGTCGCGAGCGGTTCGGGTTCGGCAGGCCCATCTCACTATAGTG 407
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
406 GGACCTGTACAGAGGGTTGGGAGCCCCCAGGAGGATGAGAGAGAGA 357
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
356 ACGTACAGAGGTTTGGGGTGATGTGAGACAGCTCATGGAGAGAGCTGAG 307
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
306 GAAGGCACCTGAGCCACAGCTCGGGGGGTAGCAGCTGACCTGACCCGCTCA 257
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
256 TCATGACCACCATGATGAGTTTTCCTCCTCATGCC 223

seq_name: gb_est2:BG064920

seq_documentation_block:
LOCUS BG064920 707 bp mRNA EST 26-JAN-2001
DEFINITION H3025D01-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3025D01 3', mRNA sequence.
ACCESSION BG064920
VERSION BG064920.1 GI:12547483
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 707)

34 lUpProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
242 AGCCACTGTGGCCCTGACCTCCAGAGCTGGCAAAACTGTGGCCTAGA 291
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
292 GGAGGTCGCGAGCGGTTCGGGTTCGGCAGGCCCATCTCACTATAGTG 341
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
342 GGACCTGTACAGAGGGTTGGGAGCCCCCAGGAGGATGAGAGAGAGA 391
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
392 ACGTACAGAGGTTTGGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 441
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
442 GAAGGCACCTGAGCCACAGCTCGGGGGGTAGCAGCTGACCTGACCCGCTCA 491
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
492 TCATGACCACCATGATGAGTTTTCCTCCTCATGCC 525

seq_name: gb_est1:AW536404

seq_documentation_block:
LOCUS AW536404 642 bp mRNA EST 31-AUG-2000
DEFINITION G0104A11-3 NIA Mouse E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0104A11 3', mRNA sequence.
ACCESSION AW536404
VERSION AW536404.1 GI:7178821
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 642)
Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: G0104 row: A column: 11
Seq primer: -21M13 Forward
High quality sequence stop: 642
POLYA-Yes.

Location/Qualifiers
1. 642
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G0104A11"
/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/note="vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an Oligo(dT) primer (NotI primer-adaptor
from GibcoBRL)
[5'-pGACTAGTCTAGTCGCGAGCGGCCGCTTTTTCCTTTT-3']
from 0.5µg of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by

AUTHORS Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka, T.S., Carter, M.G., and KO, M.S.H.
TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other_ESTs: H3025D01-5
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdnaelg@nig.nih.gov
 This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3025 row: D column: 01
 Seq primer: -21M13 forward
 High quality sequence stop: 707
 POLYA=Yes.

FEATURES

source
 1. .707
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3025D01"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryos cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with oligo(dT) -Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132. (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
 BASE COUNT 154 a 209 c 146 g 198 t
 ORIGIN

alignment_scores:
 Quality: 680.00 Length: 128
 Ratio: 5.354 Gaps: 0
 Percent Similarity: 99.219 Percent Identity: 98.438
 alignment_block:
 US-09-327-750d-30 x BG064920/rev
 Align seg 1/1 to reverse of: BG064920 from: 1 to: 707
 1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17
 606 ATGGAGTCCAAAGATCAAGGCGTGAATAATCTCAACATGGAGATGACCA 557
 17 sGlnLysLysGluGluGlyValLysProGlnAspThrIleArgArg 34
 556 TCAGAAAGAGGAGAGAGGAGAGAAACCAAGATACCATCAGAGGG 507
 34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
 506 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 457
 51 GlyGlyArgArgPheArgValArgGlnProIleAlaHistyrArgTr 67

|||||
 456 GGAGTGGCAGGCGGTTCGGGGTTCGGCAGCCCATCGCTCACTATAGATG 407
 67 PASpLeuMetGlnArgValGlyValProGlnGlyArgMetArgGluGluA 84
 406 GGAGCTGATCAGAGGTTGGGGAGCCCGAGGAGGAGGAGGAGGAGGAGA 357
 84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
 356 ACGTACAGAGGTTTGGGGTGATGTGAGACAGCTCATGGAGAAGCTGAGG 307
 101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
 306 GAAAGGAGCTGAGCCAGCGTGGCGGTTAGCAGCTGACCCGCCCTCA 257
 117 SHISAspHisAspGluPheCysLeuMetPro 128
 256 TCATGACCACCATGATGAGTTTTCCTCATGCC 223
 seq_name: gb_esf2:BG277659
 seq_documentation_block:
 LOCUS BG277659 575 bp mRNA EST 21-FEB-2001
 DEFINITION ux47c11.y1 Soares_NMAX_maxillary_process Mus musculus cDNA clone
 IMAGE:3513237 5' similar to TR:Q9RJ2 Q9RJ2 BEX1 PROTEIN. ; mRNA
 SEQUENCE.
 ACCESSION BG277659
 VERSION BG277659.1 GI:13073183
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 575)
 REFERENCE NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Other_ESTs: ux47c11.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapb-r@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (infoimage.llnl.gov) for further information.
 MGI:1393813
 Seq primer: -40RP from Gibco
 High quality sequence stop: 462.
 FEATURES
 Location/Qualifiers
 1. 575
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:3513237"
 /clone_lib="Soares_NMAX_maxillary_process"
 /tissue_type="maxillary process"
 /lab_host="DH10B (phage-resistant)"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGAAGTGGAGGCGCGCCCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 156 a 136 c 201 g 81 t 1 others
 ORIGIN

alignment_scores:
 Quality: 670.00 Length: 128
 Ratio: 5.317 Gaps: 0
 Percent Similarity: 98.438 Percent Identity: 97.656
 alignment_block:

US-09-327-750D-30 x BG277659

Align seg 1/1 to: BG277659 from: 1 to: 575

```

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17
|||||
192 ATGGAGTCCAAAGATCAAGCGGTGAAATAATCTCAACATGGAGATGACCA 241
|||||
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArg 34
|||||
242 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 291
|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
292 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 341
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
342 GGAGTGCAGCGGTTCCGGGTTCCGAGCCCATCGCTCACTATAGATG 391
|||||
67 pasLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
392 GGACTGATGCAGAGGTTGGGAGGCCCGGAGGAGGATGAGAGAGGAGA 441
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
442 ACGTACAGAGGTTGNGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 491
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
492 GAAAGGAGCTGTAGCCACACCTCGCGGGGTTAGCACTGACCCGCCCTCA 541
|||||
117 sHisAspHisAspGluPheCysLeuMetPro 128
|||||
542 TCATGACCACTGATGATGATGTTGCCTCATGCCCC 575
|||||

```

seq_name: gb_est2:BG870503

```

seq_documentation_block: 756 bp mRNA EST 29-MAY-2001
LOCUS BG870503 602791422F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922644 5',
DEFINITION mRNA sequence.
ACCESSION BG870503
VERSION BG870503.1 GI:14221043
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 756)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10842 row: 1 column: 05
High quality sequence stop: 756.
Location/Qualifiers
1..756
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4922644"
/lab_host="NCI_CGAP_SG2"
/note="organ: salivary gland; Vector: pcwv-sport6; Site:1"

```

FEATURES

```

source
1..756
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4922644"
/lab_host="NCI_CGAP_SG2"
/note="organ: salivary gland; Vector: pcwv-sport6; Site:1"

```

NotI; Site_2: Sali; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 207 a 164 c 230 g 155 t

ORIGIN

alignment_scores:
Quality: 670.00 Length: 129
Ratio: 5.276 Gaps: 1
Percent Similarity: 98.450 Percent Identity: 98.450

alignment_block:

US-09-327-750D-30 x BG870503

Align seg 1/1 to: BG870503 from: 1 to: 756

```

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17
|||||
165 ATGGAGTCCAAAGATCAAGCGGTGAAATAATCTCAACATGGAGATGACCA 214
|||||
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArg 34
|||||
215 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 264
|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
265 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 314
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
315 GGAGTGCAGCGGTTCCGGGTTCCGAGCCCATCGCTCACTATAGATG 364
|||||
67 pasLeuMetGlnArgValGlyGluProGlnGlyArgMet ArgGluGlu 83
|||||
365 GGACCTGATGCAGAGGTTGGGAGGCCCGGAGGAGGATGACGAGAGGAG 414
|||||
84 AsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuAr 100
|||||
415 AACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTGAG 464
|||||
100 gGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
465 GGAAGAGGAGCTGAGCCACACCTCGCGGGGTTAGCACTGACCCGCCCTC 514
|||||
117 sHisAspHisAspGluPheCysLeuMetPro 128
|||||
515 ATCATGACCACTGATGATGATGTTGCCTCATGCCCC 549
|||||

```

seq_name: gb_est1:BB654459

```

seq_documentation_block: 499 bp mRNA EST 06-SEP-2000
LOCUS BB654459 UI-M-AJ1-aha-f-10-0-UI.f1 NIH_BMAP_MOB_N Mus musculus cDNA clone
DEFINITION UI-M-AJ1-aha-f-10-0-UI 5', mRNA sequence.
ACCESSION BB654459
VERSION BB654459.1 GI:9980372
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 499)
AUTHORS Ronaldo M.F., Lennon G. and Soares M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706

```

Fax: 301 443 9890

Email: mestr@mail.nih.gov

cDNA Library Preparation: M.B. Soares Lab Clóne distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements

Seq primer: M13 Reverse

FEATURES

Location/Qualifiers

source

1..499

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-AJ1-aba-f-10-0-UI"

/clone_lib="NIH_BMAP_MOB_N"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_MOB_N library is a normalized library constructed from mouse olfactory bulbs. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories."

146 a 108 c 179 g 66 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 652.00 Length: 123
Ratio: 5.344 Gaps: 0
Percent Similarity: 99.187 Percent Identity: 99.187

alignment_block:

US-09-327-750D-30 x BE654459

Align seg 1/1 to: BE654459 from: 1 to: 499

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17
131 ATGGAGTCCAAAGATCAAGCGGTGAAAAATCTCAACATGGAGATGACCA 180
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleArgArg 34
181 TCAGAAAGAGGAGGAGAGAGAGAAAGCCACAGATACCATCAGAGGG 230
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
231 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 280
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
281 GGAGTTCGAGCGGTTCGGGTTCGGAGCCCATCGCTCACTATAGTG 330
67 pasPLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluA 84
331 GGACCTGATGCGAGAGGTGGGGAGCCCGGAGGAGGATGAGAGAGAGA 380
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
381 ACGTACAGAGTTGGGGGTGATGTGACACAGCTCATGGAGAACTGAGG 430
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
431 GAAAGGCAGCTGAGCCACAGCTCGGGCGGTAGCAGCTGACCGCGCTCA 480
117 sHisAspHisAspGlu 123
481 TCATGACCACTATGATGAG 499

seq_name: gb_estl:AW536974

seq_documentation_block:

LOCUS AW536974

DEFINITION G0110H10-3 NTA Mouse E7.5 Embryonic Portion cDNA Library Mus

ACCESSION AW536974

VERSION AW536974.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

REFERENCE 1 (bases 1 to 577)

AUTHORS Tanaka, T.S., Jaredat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac

M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W.H.

III, Becker, K.G. and Ko, M.S.H.

TITLE Genome-wide expression profiling of mid-gestation placenta and

JOURNAL embryo using a 15,000 mouse developmental cDNA microarray

MEDLINE Prof. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

COMMENT 20381348

Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@gsun.grc.nia.nih.gov

Plate: G0110 row: H column: 10

Seq primer: -21M13 Forward

High quality sequence stop: 577

POLYA=Yes.

FEATURES

Location/Qualifiers

1..577

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="G0110H10"

/clone_lib="NTA Mouse E7.5 Embryonic Portion cDNA Library"

/sex="unknown"

/dev_stage="7-5dpc Embryo"

/lab_host="DH10B"

/note="vector: pSPORT1 (Gibco/BRL Life Technology);

Site_1: SalI; Site_2: NotI; Total RNAs were extracted from

6 Embryo. The double-stranded cDNA was synthesized by

Gibco's kit with an Oligo(dT) primer (NotI primer-adaptor

from GibcoBRL)

[5'-pGACTAGTCTCATCGAGCGCGCCCTTTTTTTTTTTT-3']

from 0.5ug of mRNA. The double-stranded cDNAs were

treated with T4 DNA polymerase and purified by

ethanol-precipitation. The cDNAs were ligated to

Lone-linker LL-Sal3 (include SalI sequence). The cDNAs

were purified by phenol/chloroform and separated from

free linkers by Centricon 100. Then, cDNAs were amplified

by long-range high fidelity PCR using Takara's Ex Taq

polymerase. Then, the cDNAs were purified by

phenol/chloroform and by Centricon 100. The cDNAs were

digested with SalI and NotI enzymes. Then, the cDNAs were

size selected by Gibco's Size Fractionation Column. The

cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid

vector. The DH10B E. coli host was transformed with the

ligation mixture by chemical method. The library was

constructed by Xiaohong Wang and Minoru S. H. Ko."

BASE COUNT 137 a 160 c 126 g 154 t

ORIGIN

alignment_scores:

Quality: 632.00 Length: 118
Ratio: 5.402 Gaps: 0
Percent Similarity: 99.153 Percent Identity: 99.153

alignment_block:

US-09-327-750D-30 x AW536974/rev

Align seg 1/1 to reverse of: AWS36974 from: 1 to: 577

```

11 LeuAsnMetGluAsnAspHisGlnLysGluGluLysGluLysGluLysPr 27
|||||
576 CTCACATGGAGATGACCATCACAAAAGAGGAGGAGAGGAGAGAGCC 527
27 oglnAspThrIleArgArgGluProAlaValAlaLeuIleSerGluAlaG 44
|||||
526 ACAAGATACCATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 477
44 lYlYsAsnCYsAlaProArgGlyArgArgArgPheArgValArgGln 60
|||||
476 GCAAAACACTGTCACCTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 427
61 ProIleAlaHisTyrArgTrrPAsPleuMeGlnArgValGlyGluProG 77
|||||
426 CCCATCGCTCACTATAGATGGACCTGATCAGAGGAGGAGGAGGAGGAG 377
77 nGlyArgMetArgGluGluAsnValGlnArgPheGlyGlyAspValArg 94
|||||
376 GGGAGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 327
94 lNleuMeGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAla 110
|||||
326 AGCTCATGAGGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 277
111 ValSerThrAspProProHisHisAspHisAspHisAspHisAspHis 127
|||||
276 GTTAGCTAGACCCGCTCATCATGACCAACCATGATGATGATGATGAT 227
127 tPro 128
226 GCCC 223

```

seq_name: gb_est2:BF607762

```

seq_documentation_block: 785 bp mRNA EST 01-APR-2001
LOCUS BF607762
DEFINITION MT1_000704 Mouse 9-day fetus cDNA library ICRP522 Mus musculus
cDNA clone ICRP522A1947 5', mRNA sequence.
ACCESSION BF607762
VERSION BF607762.1 GI:13504254
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B.G.,
Lehrach, H. and O'Brien, J.
TITLE Detection of a high number of novel genes in a 9-day mouse embryo
cDNA library normalised by oligonucleotide fingerprinting
JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
Laboratory 123, Dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhestr. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
EST's are made from clones being representatives of clone clusters.
Clone clusters were calculated from oligonucleotide fingerprints.
PCR Primers
FORWARD: 5'-GAGTATTCCAGAGTAGTGA-3'
BACKWARD: 5'-TAATAGCTACATATAGG-3'
Seq primer: 5'-ATTAGTGCACATAG-3'
High quality sequence stop: 785.
location/Qualifiers
1. 785
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="ICRP522A1947"
/clone_lib="Mouse 9-day fetus cDNA library ICRP522"

```

FEATURES source

```

/tissue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E.coli, XL1 blue"
/notes="Vector: pSVSPORT1; Site.1: NotI; Site.2: SalI;
Library preparation by oligo.dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."
BASE COUNT 190 a 188 c 233 g 153 t 21 others
ORIGIN

```

```

alignment_scores:
Quality: 591.00 Length: 128
Ratio: 4.844 Gaps: 4
Percent Similarity: 95.312 Percent Identity: 93.750
alignment_block:
US-09-327-750d-30 x BF607762 ..
Align seg 1/1 'to: BF607762 from: 1 to: 785

```

```

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17
|||||
259 ATGGAGTCCAAAGATCAAGCGGTGAAAATCTCAACATGGAGAATGACCA 308
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleArgArg 34
|||||
309 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 358
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCYsAlaProArg 50
|||||
359 AGCCAGCTGTGCCCTGACCTCCGAGGCTGGCAAAAACCTGTGCACCTAGA 408
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
409 GGAGGTGCGAGCGGTTCGGGTTCGGCAGCCCATCGCTCACTATAGATG 458
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
459 GGACCTGATGTCAGAGGGTTGGGAGGCCCGGAGGAGGATGAGAGGAGGA 508
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMet. GluLysLeuAr 100
|||||
509 ACGTACAGAGGTTTGGGGTGATGTGACACAGETCATGGGGGAACCTGAG 558
100 gLluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp.ProPro 116
|||||
559 GGAAGGCACTGAACCCACAGCCTCGGGCGGTTAGCAGCTGACCCCGCCT 608
117 His.HisAspHisHis.AspGluPhe 124
|||||
609 CATTCATGACCACCATGATGATGTTT 634

```

seq_name: gb_est2:BG228077

```

seq_documentation_block:
LOCUS BG228077 587 bp mRNA EST 08-FEB-2001
DEFINITION ux47c11.x1 Soares.NMAMAX_maxillary process Mus musculus cDNA clone
IMAGE:3513237 3' similar to TR:09R1J2 09R1J2 BEX1 PROTEIN. ;, mRNA
sequence.
ACCESSION BG228077
VERSION BG228077.1 GI:12715592
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 587)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

```


(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 131 a 197 c 132 g 176 t 1 others
ORIGIN

alignment_scores:
Quality: 481.50 Length: 117
Ratio: 4.500 Gaps: 1
Percent Similarity: 91.453 Percent Identity: 84.615

alignment_block:

US-09-327-750D-30 x A2936393/rev

Align seg 1/1 to reverse of: A2936393 from: 1 to: 637

```

1 MetGluSerLys...AspGlnGlyValLysAsnLeuAsnMetGluAsnAs 16
|||||
382 ATGGAGTCCAAGTGGACACAGCGCTGAAANCTCAACATGAGGATGA 333
16 PhisGlnLysLysGluGluLysGluLysProGlnAspThrIleA 33
|||||
332 CCATCAGGAAAGGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 283
33 rgGluProAlaValLeuLeuSerGluAlaGlyLysAsnCysAlaPro 49
|||||
282 GGGATCCGATGTGGCCCTCGCTTCGAGCTGGAGCTACTACGTGCCT 233
50 ArgGlyArgArgArgPheArgValArgGlnProIleAlaHisTyr 66
|||||
232 AGAGGAGTGCAGCGCGTTCGGGTTCCGCGAGCCCATCGTCACTAC 183
66 gtrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArg 83
182 ATGGACCTGATGCTAGCGTTGGGAGCCCGGAGGAGGATGAGAGG 133
83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLys 99
|||||
132 AGAAGCTACAGAGGTTTGGGATGATGTGAGACAGCTCATGGAGAAG 83
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp 116
|||||
82 GG.GAAGGCACCTAGCCACAGCGCTCGCGCGGTGTAGCACTGACCC 34
116 o 116
33 T 33

```

seq_name: gb_est2.B1183527

seq_documentation_block:
LOCUS B1183527 701 bp mRNA 10-JUL-2001
DEFINITION UNL-P-FN-by-f-07-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-by-f-07-0-UNL 3', mRNA sequence.
ACCESSION B1183527
VERSION B1183527.1 GI:14657936
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 701)
Caetano, A.R., Johnson, R.K. and Pomp, D.
Generation and sequence characterization of a normalized cDNA library from swine ovarian follicles
Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized porcine ovarian follicles library
Seq primer: M13 -29
POLYA=Yes.

FEATURES
source

Location/Qualifiers

```

1..701
/organism="Sus scrofa"
/strain="University of Nebraska, Lincoln Swine Selection Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-by-f-07-0-UNL"
/clone_lib="UNL-P-FN"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN library is a normalized library representing porcine ovarian follicles, ranging between 2.0 to 10.0 mm in diameter, collected during 7 days of the follicular phase of the pig estrous cycle. This library was derived from the library UNL-P-F2. The tag is a string of 5-6 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996.
TAG_LIB=UNL-P-FN
TAG_TISSUE=porcine ovarian follicles
TAG_SEQ=CACACT"
BASE COUNT 173 a 190 c 119 g 218 t 1 others
ORIGIN

```

alignment_scores:

Quality: 480.50 Length: 130
Ratio: 4.215 Gaps: 2
Percent Similarity: 87.692 Percent Identity: 71.538

alignment_block:

US-09-327-750D-30 x B1183527/rev

Align seg 1/1 to reverse of: B1183527 from: 1 to: 701

```

1 MetGluSerLysAsp.....GlnGlyValLysAsnLeuAsnMetGluAs 15
|||||
653 ATGGCGTCCAAAGAGGAGGAGCAGCAAGCAAGTGAATAATCTCAGCATG 604
15 nAspHisGlnLysGluGluLysGluLysProGlnAspThrIleA 32
|||||
603 TACCACCAAGGAGAAATGAAGAAAGAGAAAGGCGGCAAGATGCTAATA 554
32 rgArgGluProAlaValAlaLeuLeuSerGluAlaGlyLysAsnCysAla 48
|||||
553 AAGGAGAGCGCT...TTGGCCCTCCCTGTGGAAGCTGGTGATAATGGCTA 507
49 ProArgGlyArgArgArgPheArgValArgGlnProIleAlaHisTyr 65

```

```

|||||.....|
506 CCTAGAGGAATCTAGCGGTTCCGTGTGAGGAGCCCTATCTCTGCAGTA 457
65 rArgTTPasPleuMetGlnArgValGlyGluProGlnGlyArgMetArg 82
|||||.....|
456 TAGATGGGATATAGCAGGAGGCTTGGAGAGCCACAGGAGATGAGAG 407
82 luGluAsnValGlnArgPheGlyClyAspValArgGlnLeuMetGluLys 98
|||||.....|
406 AAGAGAATATGAAAGGATTGGGAGAGGAGGCTGCTGATGGAAGAAG 357
99 LeuArgGluArgGlnLeuArgHisSerLeuArgAlaValSerThrAspPr 115
|||||.....|
356 CTGAGGGAAGAGAGTGTAGTCATGTCTCGGGCAGTTAGCACATGACC 307
115 pProHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||.....|
306 CCTCACCATGACCATCAGCATGATGATTTTGCCTTATGCCT 267

```

seq_name: gb_est2:D76670

seq_documentation_block: 370 bp mRNA EST 07-OCT-1996
LOCUS D76670
DEFINITION MUS77D10 mouse embryonal carcinoma cell line F9 Mus musculus cDNA
clone 77D10, mRNA sequence.

ACCESSION D76670

VERSION D76670

KEYWORDS EST

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 370)

AUTHORS Jishiguchi,S., Sakuma,R., Nomura,M., Zou,Z., Jearnaisilavong,J.,

Toh,T., Yasunaga,T. and Shimada,K.

TITLE A catalogue of genes in mouse embryonal carcinoma F9 cells

identified with expressed sequence tags

J. Biochem. (Tokyo) 119 (4), 749-767 (1996)

J. Biochem. (Tokyo) 119 (4), 749-767 (1996)

COMMENT Contact: Kazunori Shimada

Department of Medical Genetics, Division of Molecular Biomedicine

Research Institute for Microbial Diseases, Osaka University

3-1, Yamadaoka, Suita, Osaka, 565, Japan

Tel: 06-879-8325

Fax: 06-879-8326.

FEATURES Location/Qualifiers

source

1..370

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="77D10"

/clone_lib="mouse embryonal carcinoma cell line F9"

BASE COUNT 110 a 72 c 124 g 53 t 1 others

ORIGIN

alignment_scores:

Quality: 476.50 Length: 96

Ratio: 5.124 Gaps: 1

Percent similarity: 96.875 Percent identity: 96.875

alignment_block:

US-09-327-750D-30 x D76670

Align seg 1/1 to: D76670 from: 1 to: 370

1 MetGluSerTyrAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH 17

68 ATGGAGTCCAAAGATCAAGCGGTGAAATCTCAATGAGATGACCA 117

17 scLysLysGluGluLysGluLysProGlnAspThrLeuArgArg 34

118 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 167

```

34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||.....|
168 ACCCAGCTGTGCCCTGATCTCCGAGGCTGGCAAAACTGTGGCCCTAGA 217
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||.....|
218 GGAGGTCCGAGCGGTTCCGGGTTCCGGCAGCCCATCGCTCACTATAGATG 267
67 pAsPleuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||.....|
268 GGACCTGATGCAGAGGTTGGGAGCCCCAGGAGGATGAGAGGAGA 317
84 snValGlnArgPheGlyCly...AspValArgGlnLeu 95
|||||.....|
318 ACGTACAGAGGTTTNGGGGTTATTGTTAGCAACTT 355
seq_name: gb_hlc:AK003429

```

seq_documentation_block:

LOCUS AK003429

DEFINITION Mus musculus 18 days embryo cDNA, RIKEN full-length enriched

library, clone:1110004J10, full insert sequence.

ACCESSION AK003429

VERSION AK003429.1

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA,

clone:1110004J10.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 480)

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

Methods in enzymology. 303, 19-44 (1999)

JOURNAL 99279253

MEDLINE 10349636

PUBMED 2 (bases 1 to 480)

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome research. 10 (10), 1617-1630 (2000)

JOURNAL 20499374

MEDLINE 11042159

PUBMED 3 (bases 1 to 480)

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kitsumi,T., Tashiro,H., Itoh,M.,

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Nishiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome research. 10 (11), 1757-1771 (2000)

JOURNAL 20530913

MEDLINE 11076861

PUBMED 4 (bases 1 to 480)

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

JOURNAL 20530913

MEDLINE 11076861

PUBMED 5 (bases 1 to 480)

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Araki,A.,

Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,

Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,

Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,

Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,

Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,

Okido,T., Owa,C., Saito,H., Saito,K., Sakai,K., Sano,H.,

Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

OM of: US-09-327-750d-30 to: Issued_Patents_NA.* out_format : pfs

Date: Mar 11, 2002 3:50 PM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-frame-p2n.model -DEV-xlh
-Q/cgn2_1/USPTO.spool/US09327750/runat_11032002_101154_20340/app_query.fasta_1.1472
-DB-Issued_Patents_NA -OPMT-Fastcap -SUFFIX-rni -GAPOP-12.000
-GAPEXT-4.000 -MINMATCH-0.100 -LOOPEXT-0.000 -LOOPEXT-0.000
-GAPOP-4.500 -GAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500
-GAPOP-6.000 -GAPEXT-7.000 -XGAPOP-10.000 -XGAPEXT-0.500
-DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-blosum62
-TRANS-human40.cdi -LIST-45 -DOCALIGN-200 -THR_SCORE-PCT
-THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs
-NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09327750/cgn1_1.99 -NCPU-6 -TCPU-3 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -NO_XLPXY -WAIT -THREADS-1

Search information block:

Query: US-09-327-750d-30

Query length: 128

Database: Issued_Patents_NA.*

Database sequences: 351203

Database length: 113238999

Search time (sec): 146.090000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-381-881-1 +		93.50	169.00	0.0919	1315
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-281-221-1 +		93.50	169.00	0.0919	1315
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-931-999-4 +		86.50	137.69	5.09	6755
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-809-740A-1 -		85.00	142.20	2.86	3252
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-809-740A-4 -		85.00	142.20	2.86	3252
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-426-599B-3 +		83.50	156.25	0.4714	624
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-506-553C-6 +		82.50	154.49	0.5910	607
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-995-1 +		82.50	146.08	1.74	1365
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-403-852D-1 +		81.50	129.79	14.04	5392
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-510-646B-1 +		81.50	129.79	14.04	5392
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-231-818-1 +		81.00	129.79	14.04	5392
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-948-705-1 +		81.00	141.68	3.05	1552
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-393-985-17 +		81.00	138.85	4.39	2040
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-951-715A-6 +		81.00	132.89	9.44	3624
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-459-448A-6 +		81.00	132.89	9.44	3624
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-595A-6 +		81.00	132.89	9.44	3624
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-504B-6 +		81.00	132.89	9.44	3624
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-053-549-7 +		81.00	132.89	9.44	3624
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-717-294-42 +		79.50	127.68	18.40	4451
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-053-549-1 +		79.00	119.52	52.37	8854
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-147-777-3 +		78.50	107.74	237.26	24979
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-452-872-3 +		78.50	107.74	237.26	24979
/cgn2_6/ptodata/2/ina/PTUS_COMB.seq:PCT-US93-03395-3		78.50	107.74	237.26	24979
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-660-645A-4 -		77.50	140.07	3.76	908
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-298-718-4 -		77.50	140.07	3.76	908
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-546-969-4 -		77.50	140.07	3.76	908
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-980-832-1 +		77.50	116.72	75.00	8625
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-980-832-27 +		77.50	113.98	106.58	11233
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-1 +		77.00	97.60	870.99	49377
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-452-075-1 -		76.50	126.19	22.27	2841
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-231-081-1 -		76.50	108.49	215.70	15664
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-282-3 +		76.50	108.49	215.70	15664
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-066-3 +		76.50	108.49	215.70	15664
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-068-3 +		76.50	108.49	215.70	15664
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-592-874-1 -		76.50	102.17	484.98	28804
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-096-942-2 -		76.50	102.17	484.98	28804
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-096-867-2 -		76.50	102.17	484.98	28804
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-258-261B-6 +		76.50	102.12	488.43	28958

/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-456-837-6 + 76.50 102.12 488.43 28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-342-6 + 76.50 102.12 488.43 28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-646A-6 + 76.50 102.12 488.43 28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-076A-6 + 76.50 102.12 488.43 28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-4 + 76.50 102.12 488.43 28958

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-381-881-1

seq_documentation_block:
; Sequence 1, Application US/08381881
; Patent No. 5939064
; GENERAL INFORMATION:
; APPLICANT: SAVALKOU, PAUL H.,M.
; APPLICANT: GAASTRA, WILLEM
; TITLE OF INVENTION: BORDETELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,881
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TPP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; TELEX: 44-0704 "STEVENS"
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bordetella bronchiseptica
; STRAIN: 401
; IMMEDIATE SOURCE:
; CLONE: E coli PC2495(pivb3-420)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 540..1142
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1143..1315
; US-08-381-881-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.908 Gaps: 4
Percent Similarity: 51.042 Percent Identity: 33.333
alignment_block:
US-09-327-750D-30 x US-08-381-881-1

```

Align seg 1/1 to: US-08-381-881-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyGlyArgArgPheArgValar 59
542 GCAAGCCAAACAGTCTCTCTGGCGCGCGCTCGCGCGCTCGCGCTCG 591
59 gGlnProLe.....AlaHisTyrArgTrrp..... 67
592 CGGCCCATCGCGAGACGCGACCATTCATTACCGCGCAGCATCACGAC 641
68 ..AspLeuMetGlnArgValGlyGluPro.....GlnGlyArgMet 80
642 CAGACCTGCAGATCGAGGACCGCGCGCTTACATCAAGTGTGTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetG1 97
692 CTGCGCCAC.....GATCTCCAAGAGCGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
721 AGAAGCGCGCGAGCTGGCGGCGCGACTCGCTTCGATATCAAGCTGAAG 770
114 spProProHisHisAspHisAspGluPheCysLeu 126
771 GACTGCGCGACCGCTCAACACTCTCAAGCTGTACTT 808
seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-09-281-221-1
seq_documentation_block:
; Sequence 1, Application US/09281221
; Patent No. 6284256
; GENERAL INFORMATION:
; APPLICANT: GASTRA, WILLEM
; TITLE OF INVENTION: BORDETTELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/281,221
; FILING DATE: 30-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/381,881
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TPP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```

```

ORIGINAL SOURCE:
ORGANISM: Bordetella bronchiseptica
STRAIN: 401
IMMEDIATE SOURCE:
CLONE: E coli PC2495 (pIVB3-420)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..539
FEATURE:
NAME/KEY: CDS
LOCATION: 540..1142
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1143..1315
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-281-221-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.908 Gaps: 4
Percent Similarity: 51.042 Percent Identity: 33.333

alignment_block:
US-09-327-750D-30 x US-09-281-221-1 ..
Align seg 1/1 to: US-09-281-221-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyGlyArgArgPheArgValar 59
|||||:||||| ||| ||| ||| ||| |||:|||||
542 GCAAGCCAAACAGTCTCTCTGGCGCGCGCTCGCGCGCTCGCGCTCG 591
59 gGlnProLe.....AlaHisTyrArgTrrp..... 67
|:||||| :|||||
592 CGGCCCATCGCGAGACGCGACCATTCATTACCGCGCAGCATCACGAC 641
68 ..AspLeuMetGlnArgValGlyGluPro.....GlnGlyArgMet 80
||||| :||||| ||| ||| ||| |||:|||||
642 CAGACCTGCAGATCGAGGACCGCGCGCTTACATCAAGTGTGTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetG1 97
||| :||| |||:|||||
692 CTGCGCCAC.....GATCTCCAAGAGCGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
|||: |||:||||| ||| ||| ||| ||| |||: |||
721 AGAAGCGCGCGAGCTGGCGGCGCGACTCGCTTCGATATCAAGCTGAAG 770
114 spProProHisHisAspHisAspGluPheCysLeu 126
||| ||| :||||| |||
771 GACTGCGCGACCGCTCAACACTCTCAAGCTGTACTT 808
seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-931-999-4
seq_documentation_block:
; Sequence 4, Application US/08931999
; Patent No. 6043219
; GENERAL INFORMATION:
; APPLICANT: Iandolo, John J.
; APPLICANT: Crupper, Scott S.
; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```


OTHER INFORMATION: /note= "ORF2"
UIS-08-809-740A-4

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,599B
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 1604-123A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-624-1589
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Optimized cDNA"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..621
; US-08-426-599B-3

alignment_scores:
    Quality: 83.50      Length: 96
    Ratio: 1.670        Gaps: 6
    Percent Similarity: 52.083      Percent Identity: 34.375

alignment_block:
    US-09-327-750D-30 x US-08-426-599B-3 ..

Align seg 1/1 to: US-08-426-599B-3 from: 1 to: 624
33 ArgGluProAlaValAlaLeuLeuSerGluAlaGlyLysAsnCysAlaApr 49
|||||
15 CGTCAACG.....CTGCTCTACACCGCTGGCGCGCGCCCTCCAA... 56
49 oArgGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrA 66
|||||
57 .....GGGCGACGCGCCCTCGACGCGCGCGCGCGCGCACTCACTCC 99
66 rGtrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGlu 82
|||||
100 GG.....CGAGTCTGGACCATCAACGTCGAGCGCGCGCAC 134
83 GluAsnValGlnArgPheGlyArgValArgGlnLeuMetGluLysLe 99
|||||
135 CAAGCGCGCAAGATCTGGCGCGCAC...CGACTGCTACTTCGACGACT 181
99 uArgGluArgGlnLeuSerHis.....CGCGCGCGCGCGCGCGCG 106
|||||
182 CGGCGCGCGCAATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 231
107 ..SerLeuArgAlaValSerThrAspProProHisHis 118
|||||
232 AAGCGCTCGG.....CGCGCGCGCGCGCGCGCGCGCGCGCGCG 257

seq_name: /cqn2.6/ptodata/2/1na/6A_COMB.seq.us-08-506-553C-6

seq_documentation_block:
; Sequence 6, Application US/08506553C
; Patent No. 6120989
; GENERAL INFORMATION:
; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Sonneborn, Han-H.;
; APPLICANT: Plachter, Bodo; and Jahn, Gerhard
; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF

```

```

; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/506,553C
; FILING DATE: 07/25/95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LEADER 202-PFF/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-506-553C-6

alignment_scores:
    Quality: 82.50      Length: 117
    Ratio: 1.650        Gaps: 4
    Percent Similarity: 42.735      Percent Identity: 26.496

alignment_block:
    US-09-327-750D-30 x US-08-506-553C-6 ..

Align seg 1/1 to: US-08-506-553C-6 from: 1 to: 607
50 ArgGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrA 66
|||||
107 CGGCTCGGAGGAGGTGGCGCGGTGGCGCGCAAGAA...GCACGACCGG 153
66 gTrpAspLeuMetGlnArgValGlyGluProGln..... 77
|||||
154 GTGCGCGCGGTGGTTCGGTACGCGGAAATGAGTAGCGGTGGCGCGGC 203
77 ..... 77
204 GGTATCATGACACGGTCTTTCTCCAGGAATAATACGACGACCA 253
78 .....GlyArgMetArgGluG1 83
|||||
254 GATCACCAGCTACTGACGTCCAAAGGTGGATCGGCGCGCGGAGGAG 303
83 uAsnValGlnArgPheGlyArgValArgGlnLeuMetGluLysLeuA 100
|||||
304 GAGGAGCGCGGTGGATCGCACTCGGCAATCTTCAACGACCGG 353
100 rG.....GluArgGlnLeuSer..... 105
|||||
354 AAAGAGAGAGCGACAGCGAGGATCTGTAACTGAGTTTCGTCCTCC 403
106 HisSerLeuArgAlaValSerThrAspProProHisHisAspHis 122
|||||

```

404 CACCAAGAGCAAAAGTCGGGAGATCCTGAAGAGCAGCAGCGGCATGA 453

122 p 122

454 A 454

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-030-995-1

seq_documentation_block:

: Sequence 1, Application US/09030995

: Patent No. 6235519

: GENERAL INFORMATION:

: APPLICANT: Wang, Yongzhao

: APPLICANT: Childs, John D.

: APPLICANT: Squires, Charles H.

: TITLE OF INVENTION: A Gene Involved in Thiophene

: TITLE OF INVENTION: Biotransformation from No. 6235519ardia Asteroides KGB1

: FILE REFERENCE: EBC97-05

: CURRENT APPLICATION NUMBER: US/09/030,995

: CURRENT FILING DATE: 1998-02-26

: NUMBER OF SEQ ID NOS: 3

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 1

: LENGTH: 1365

: TYPE: DNA

: ORGANISM: No. 6235519ardia asteroides

: US-09-030-995-1

alignment_scores:

Quality: 82.50

Ratio: 1.231

Percent Similarity: 54.472

Length: 123

Gaps: 8

Percent Identity: 30.894

alignment_block:

US-09-327-750D-30 x US-09-030-995-1

Align seg 1/1 to: US-09-030-995-1 from: 1 to: 1365

27 ProGlnAspThrIleArgArg...GluProAlaValAlaLeuIleSerG1 42

|||||:|||||:|||||: ||| :|||||: |||

908 CTTATGTCACAAATCCGCGCTGTTCATCGCATCGCAGGACCATGG 957

42 uAlaGlyLysAsnCysAla...ProArgGlyValArgArg... 55

|||: |||: |||: |||: |||: |||

958 TTCACCTTCACAGTTCGAGCGCCAGCGCTGTGGCGCGCAGCTGAT 1007

56PheArgValArgGlnProIleAlaHisTyrArgTrpAspLeuMet 70

|||: |||: |||: |||: |||: |||

1008 CTGGCGCGCATCGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGATG 1057

71 GlnArgValGlyGlu.....Pro.GlnGlyArgMetArgGluG 83

|||||: |||: |||: |||: |||: |||

1058 TCGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1107

83 luAsnValGln.....ArgPheGlyAspValArgGlnLeuMet 96

|||: |||: |||: |||: |||: |||

1108 ATCCGCTATCGGGCGAGTACGTACGTCAAGGAGCTGATCGCGAGCGACTA 1157

97 GluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerTh 113

|||||: |||: |||: |||: |||: |||

1158 TCGAGCTTCGAGCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1207

113 rAsp.....ProPro..... 116

1208 AGCACAAGAGAAGAACATCATGAGTTCGCGCCACCAACTCCTATCGCTCG 1257

117 ..HisHisAspHisHis 121

1258 GTCATCAGCGCGCACCAT 1274

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-403-852D-1

seq_documentation_block:

: Sequence 1, Application US/08403852D

: Patent No. 5891695

: GENERAL INFORMATION:

: APPLICANT: Blanc, Veronique

: APPLICANT: Blanchet, Francis

: APPLICANT: Crouzet, Joel

: APPLICANT: Jacques, Nathalie

: APPLICANT: Lacroix, Patricia

: APPLICANT: Thibaut, Denis

: APPLICANT: Zagorec, Monique

: APPLICANT: Debussche, Laurent

: APPLICANT: De Crecy-Lagard, Valerie

: TITLE OF INVENTION: Polypeptides Involved In The

: TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences

: TITLE OF INVENTION: Coding For These Polypeptides And Their Use

: NUMBER OF SEQUENCES: 43

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner

: STREET: 1300 I Street, N.W., Suite 700

: CITY: Washington

: STATE: D.C.

: COUNTRY: USA

: ZIP: 20005-3315

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: OPERATING SYSTEM: IBM PC compatible

: SOFTWARE: Patent In Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/403,852D

: FILING DATE: 10-MAY-1995

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: PCT/FR 93/00923

: FILING DATE: 25-SEP-1993

: PRIOR APPLICATION DATA: FR 92/11441

: FILING DATE: 25-SEP-1992

: ATTORNEY/AGENT INFORMATION:

: NAME: Meyers, Kenneth J.

: REGISTRATION NUMBER: 25,146

: REFERENCE/DOCKET NUMBER: 03806.0054-00000

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (202) 408-4000

: TELEFAX: (202) 408-4400

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 5392 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: linear

: MOLECULE TYPE: CDNA

: HYPOTHETICAL: NO

: ANTI-SENSE: NO

: ORIGINAL SOURCE:

: ORGANISM: S.pristinaespiralis

: US-08-403-852D-1

alignment_scores:

Quality: 81.50

Ratio: 1.430

Percent Similarity: 48.305

Length: 118

Gaps: 5

Percent Identity: 27.966

alignment_block:

US-09-327-750D-30 x US-08-403-852D-1

Align seg 1/1 to: US-08-403-852D-1 from: 1 to: 5392

17 HisGlnLysLysGluGluLysGluLysProGlnAspThrIleArgAr 33

|||||: |||: |||: |||: |||: |||

4800 CATCGCTGTACCGGTTCATCGCGCGCGCGCGCGCGCGCGCGCGCG 4849


```

; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
; US-07-951-715A-6

alignment_scores:
  Quality: 81.00 Length: 121
  Ratio: 1.397 Gaps: 4
  Percent Similarity: 47.934 Percent Identity: 23.967

alignment_block:
  US-07-327-750D-30 x US-07-951-715A-6
  Al'gn seg 1/1 to: US-07-951-715A-6 from: 1 to: 3624
  14 GluAspHisGlnValProAlaValAlaLeuIleSerGluAlaGlyLysProGlnAsp 30
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1026 GAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG
  30 rIleArgArgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsn 47
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1076 TGGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1125
  47 ySala.....ProArgGlyValArgArg 55
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1126 CTGCTGTGGGGCATCTACTGGACCCCATCCGCGCGTGCACCGCGG 1175
  56 PheArgValArgGlnProIleAlaHisTyrArgTyrPaspLeuMetGln 72
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1176 CTTCACTTCACCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCA
  72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPhe 89
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1214 GCACGCCCACTACAGCCGCGCTACGAGAGCC 1247
  89 lYcLysValArgGlnLeuMetGlnLysLeuArgGluArg..... 102
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1248 .....CGGCTTCACCTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
  103 .....GlnLeuSerHisSerLeuAlaValSerThrAspProPhe 117
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1287 CGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG
  117 sHisAspHisHis 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  1337 ACATCGGCATCAT 1349
```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-459-448A-6

```

seq_documentation_block:
; Sequence 6, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,448A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
; US-08-459-448A-6

alignment_scores:
```


OM of: US-09-327-750D-31 to: EST.* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-frame1.p2n.model -DEV-xlh
-Q/cgnt2/_USPTO.spool/US09327750/runat_11032002.101153_20308/app_query.fasta.1.1472
-DB-EST -QFWT-fastap -SUFFIX-rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOAPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750_ECGN1.1.5654
-NCPU=6 -TCPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-31

Query length: 128

Database: EST.*

Database sequences: 11351937

Database length: 107921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Document
gb_est1:BE314909	+	679.00	1289.59	1.1e-62	594
gb_est1:AL520931	-	679.00	1287.82	1.4e-62	717
gb_est1:BE889313	+	679.00	1287.12	1.6e-62	772
gb_est1:BG709503	+	679.00	1286.67	1.7e-62	810
gb_est1:BG707398	+	679.00	1286.38	1.7e-62	835
gb_est1:BF038787	+	679.00	1284.29	2.3e-62	1042
gb_est1:BG385825	+	674.00	1275.65	6.9e-62	946
gb_est1:BG714374	+	672.00	1274.33	8.1e-62	736
gb_est1:BF967675	+	672.00	1273.49	9.0e-62	793
gb_est1:BF967675	+	672.00	1273.21	9.4e-62	817
gb_est1:BG707334	+	668.00	1265.62	2.5e-61	813
gb_est1:BG705843	+	666.00	1266.17	2.3e-61	511
gb_est1:AI291270	+	664.00	1259.81	5.2e-61	669
gb_est1:AI291126	+	664.00	1259.24	5.6e-61	711
gb_est1:BF373433	+	663.00	1258.77	6.0e-61	610
gb_est1:BE366012	+	663.00	1251.32	1.6e-60	897
gb_est1:BE373321	+	663.00	1251.32	1.6e-60	897
gb_est1:BF967437	+	658.00	1247.04	2.7e-60	769
gb_est1:BG715745	+	658.00	1246.48	2.9e-60	816
gb_est1:BF685596	+	657.00	1244.55	3.7e-60	818
gb_est1:BG709294	+	646.00	1223.96	5.2e-59	782
gb_est1:AL520932	+	644.00	1219.53	9.2e-59	834
gb_est1:BG715353	+	638.00	1209.80	3.2e-58	694
gb_est1:BG715353	+	636.00	1204.56	6.2e-58	807
gb_est1:BG334975	+	606.50	1151.87	5.4e-55	547
gb_est1:BG655001	+	606.50	1150.48	6.4e-55	634
gb_est1:BG472167	+	606.50	1150.13	6.7e-55	658
gb_est1:BG711303	+	606.50	1149.91	6.9e-55	674
gb_est1:BG714823	+	606.50	1149.68	7.1e-55	690
gb_est1:AV702285	+	606.50	1149.67	7.1e-55	691
gb_est1:AV705808	+	606.50	1149.67	7.1e-55	691
gb_est1:AL526209	+	606.50	1149.62	7.2e-55	699
gb_est1:BG820179	+	606.50	1149.56	7.2e-55	699
gb_est1:BE790774	+	606.50	1148.49	8.3e-55	783
gb_est1:BE790774	+	606.50	1148.36	8.4e-55	794
gb_est1:AL526247	+	603.50	1138.01	1.5e-54	683
gb_est1:AV702643	+	599.50	1138.01	3.2e-54	576
gb_est1:AV7029703	+	599.50	1135.73	4.3e-54	734
gb_est1:BG715659	+	596.00	1135.10	4.6e-54	386
gb_est1:AV732454	+	596.00	1126.82	1.3e-53	929
gb_est1:BF316327	+	592.00	1123.54	2.0e-53	585
gb_est1:AI141583	+	590.50	1117.40	4.5e-53	828
gb_est1:BG709325	+	590.50	1117.40	4.5e-53	828

gb_est1:AI826617 - 590.00 1119.74 3.3e-53 584 | AI826617 WK3503.x1 NCI_CGAF
gb_est1:AW160832 + 584.50 1110.13 1.1e-52 531 | AW160832 au76b12.y1 Schneide
gb_est1:BF346478 + 583.50 1106.22 1.9e-52 656 | BF346478 602020342F1 NCI_CG
gb_est1:AV727777 + 580.50 1101.83 3.3e-52 569 | AV727777 AV727777 HTC Homo
gb_est1:AW953922 + 580.50 1101.23 3.6e-52 607 | AW953922 EST365887 MAGE res
seq_name: gb_est1:BE314909

seq_documentation_block:
LOCUS BE314909 594 bp mRNA EST 26-OCT-2000
DEFINITION 601140074F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3049290 5',
mRNA sequence.
ACCESSION BE314909
VERSION BE314909.1 GI:9145006
KEYWORDS EST.
SOURCE human;
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM93 row: c column: 19
High quality sequence stop: 594.
Location/Qualifiers
1..594
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3049290"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pORF7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GCCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 161 a 126 c 181 g 126 t
ORIGIN

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x BE314909

Align seg 1/1 to: BE314909 from: 1 to: 594

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
141 ATGAGGTCACCAAGACGAGCGTTAAACATCTCATCGTGGAAATGT 190
|||||
17 LasGluGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
191 CAACAGGAGAAATGATGAAAAAGATGAAAGGACGAGTTCGCTAATAAG 240
|||||
34 LyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||

Align seg 1/1 to reverse of: AL320931 from: 1 to: 717

```
1 MetGluSerIysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
530 ATGGAGTCCAAAGAGGAACAGCGTTAAACAATCTCATCGTGGAAAAATGT 481
|||||
17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
480 CAACAGGAAAAATCATGAAAAAGATGAAAAGAGCAGCAAGTGTCTAATAAAG 431
|||||
34 LYGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
430 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 381
|||||
51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
380 GGAACACCGTAGGCGGGTTCGCGGTTAGCAGGCCACCTCTGCAGTATAGATG 331
|||||
67 pAspIleMethHisArgLeuGlyGluProGlnAlaArgMetArgGluAla 84
330 GGACATAATGCATAGGCGTTGGAGAGCCACAGCAAGGATGAGAGAGGAGA 281
|||||
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
280 ATATGGAAAGGATTTGGGAGGAGGTGAGCAGCACTGATGGAAAAGCTGAGG 231
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
230 GAAAAGCAGTTGATGCATATCTCGCGGCAGTCAGCACTGATGCCCTCA 181
|||||
117 shisAspHisAspGluPheCysLeuMetPro 128
180 CCATGACCATTACAGATGAGTTTGGCTTATGCC 147
|||||
```

seq_documentation_block:			
LOCUS	AL520931	717 bp	mRNA EST
DEFINITION	AL520931 LTR_NFL004_NBC2		Homo sapiens cDNA clone CS0DB002YA15 3 prime mRNA sequence.

KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens
human.
Mammalia; Eutharia; Primates; Catarrhini; Hominoidea; Hominidae; Homo.
1 (bases 1 to 717)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

1. //1/
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS08002YA15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

```
alignment_scores:
  Quality: 679.00
  Ratio: 5.305
  Percent Similarity: 100.000
  Percent Identity: 100.000
  Gaps: 0
  Length: 128
```

US-09-327-750D-31 x AL520931/rev

http://image.llnl.gov
Plate: LLAM10683 Row: a Column: 22
High quality sequence stop: 770.
Location/Qualifiers
1. 810
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4797213"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcggg); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 206 a 168 c 247 g 187 t 2 others
ORIGIN

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x BG709503
Align seg 1/1 to: BG709503 from: 1 to: 810

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
125 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCTGGAAATGT 174
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
175 CAACGAGGAAATGATGAAAGATGAAAGGAGGAGGAGGAGGAGGAGGAG 224
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
225 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTACTGTGCCTAGA 274
51 GluAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
275 GGAACCGTAGGCGGTTCCCGGTTAGGAGGAGGAGGAGGAGGAGGAGGAG 324
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGluA 84
325 GGACATAATGCTAGGCTTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 374
84 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
375 ATATGGAAGGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 424
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
425 GAAAGCAGTTCAGTCATAGTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 474
117 sHisAspHisAspGluPheCysLeuMetPro 128
475 CCATGACCATCAGATGAGTTTGCCTTATGCC 508

seq_name: gb_est2.BG709503

seq_documentation_block:
LOCUS BG709503 810 bp mRNA EST 07-MAY-2001
DEFINITION 602674781F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797213 5', mRNA sequence.
ACCESSION BG709503
VERSION BG709503.1 GI:13987904
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 810)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cspas-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

seq_name: gb_est2.BG707398
seq_documentation_block:
LOCUS BG707398 835 bp mRNA EST 07-MAY-2001
DEFINITION 602672810F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795347 5', mRNA sequence.
ACCESSION BG707398

```

VERSION      BG707398.1 GI:13983707
KEYWORDS     EST.
SOURCE       Human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 835)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
             Email: cgaabs@remail.nih.gov
             Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
             Toshlyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLNL0678 row: d column: 04
             High quality sequence stop: 814.
FEATURES     source
             1..835
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="NIH_MGC_96"
             /tissue_type="hypothalamus"
             /lab_host="DH10B"
             /note="Organ: brain; Vector: pBluescriptR (modified
             pluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
             ); Oligo-dt primed using primer 5'-TTTTTTTTTTTNN-3',
             size-selected for average insert size 2.3 kb and
             normalized to 5. This is a primary library enriched
             for full-length clones and constructed using the
             Cap-trapper method (Carninci, in preparation). Library
             constructed by M. Brownstein (NHGRI/NIH), National
             Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT   220 a 176 c 246 g 193 t
ORIGIN
alignment_scores
  Quality: 679.00 Length: 128
  Ratio: 5.305 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block
  US-09-327-750d-31 x BG707398
  Align seg 1/1 to: BG707398 from: 1 to: 835
  1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
  |||
  248 ATGGAGTCCAAAGAGGACGAGCGCTTAAACAATCTCATCTGCGAAATGT 297
  |||
  17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
  |||
  298 CAACGAGCAAAATGATGAAAGATGAAAGGAGCAAGTTGCTAATAAG 347
  |||
  34 lYGlupProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
  |||
  348 GGGAGCCCTTGCCCTACCTTGAATGTTAGTGAATACTGTGTGCCTAGA 397
  |||
  51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
  |||
  398 GGAACACCTAGGCGTTCCCGTTAGGACGCCCATCTCGACATATAGTG 447
  |||
  67 paspleMethHisArgLeuGluGluProGlnAlaArgMetArgGluGluA 84
  |||
  448 GGACATATGTCATAGGTTGGAGGAGCCACAGGCAAGGATGACAGAGGAGA 497
  |||
84 smMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
498 ATATGGAAGAGGATTGGGGAGGAGGTGACACAGCTGATGGAAGAGCTGAGG 547
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
548 GAAAGACAGTTAGTCATAGTCTGCGGCGAGTCAGCAGCTGATCCCCCTCA 597
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
598 CCATGACCATCAGCATGAGTTTTCCTTATGCC 631
|||||
seq_name: gb_est2:BF038787
seq_documentation_block:
LOCUS      BF038787 1042 bp mRNA 20-OCT-2000
DEFINITION 601462119F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3865348 5',
            mRNA sequence.
ACCESSION  BF038787
VERSION    BF038787.1 GI:10745987
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1042)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgaabs@remail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            cDNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9608 row: f column: 05
            High quality sequence stop: 732.
FEATURES   source
            1..1042
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="NIH_MGC_66"
            /tissue_type="adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: ovary; Vector: pCMV-SpORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.8 kb. Library constructed by Life
            Technologies."
BASE COUNT 308 a 200 c 296 g 237 t 1 others
ORIGIN
alignment_scores
  Quality: 679.00 Length: 128
  Ratio: 5.305 Gaps: 0
  Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block
  US-09-327-750d-31 x BF038787
  Align seg 1/1 to: BF038787 from: 1 to: 1042
  1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
  |||
  85 ATGGAGTCCAAAGAGGACGAGCGCTTAAACAATCTCATCTGCGAAATGT 134
  |||
  17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
  |||
  135 CAACGAGCAAAATGATGAAAGAGATGAAAGGAGCAAGTTGCTAATAAG 184
  |||

```

Percent similarity: 100.000 Percent Identity: 99.219

alignment_block:
US-09-327-750D-31.x BG285825

Align seg 1/1 to: BG285825 from: 1 to: 946

1 MetGluSerLySgLUgLUaArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
185 GGGAGCCCTGGCCCTACTTGAATGTTAGTGAATCTCTGTCCTGCTAGA 234
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyArgTr 67
235 GGAACCGTAGCGGTTCCGGCTTAGGCAGCCCATCTCGCAGTATAGATG 284
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
285 GGACATAATGCTAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 334
84 snMetGluArgIleGlyGluValArgGlnLeuMetGluLyLeuArg 100
335 ATATGAAAGGATCGGGGAGGAGGTGACACAGCTGATGGAAGAGCTGAGG 384
101 GluLySgLUgLUaSerHisSerLeuArgAlaValSerThrAspProHI 117
385 GAAAGCAGTTGAGTCATAGTTTGGCGGCGAGTCAGCAGCTGATCCCTCA 434
117 sHisAspHisArgGluPheCysLeuMetPro 128
435 CCATGACCATCAGCATGAGTTTGGCTTATGCC 468

seq_name: gb_est2:BG285825

seq_documentation_block:
LOCUS BG285825 946 bp mRNA EST 21-FEB-2001
DEFINITION 602380917F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:449845 5',
mRNA sequence.
ACCESSION BG285825
VERSION BG285825.1 GI:13038171
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10361 row: a column: 06
High quality sequence stop: 710.
Location/Qualifiers
1..946
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:449845"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 276 a 201 c 279 g 190 t
ORIGIN

alignment_scores:
Quality: 674.00 Length: 128
Ratio: 5.266 Gaps: 0

seq_name: gb_est2:BG714974

seq_documentation_block:
LOCUS BG714974 726 bp mRNA EST 08-MAY-2001
DEFINITION 602677263F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799823 5',
mRNA sequence.
ACCESSION BG714974
VERSION BG714974.1 GI:13993905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10689 row: n column: 16
High quality sequence stop: 723.
Location/Qualifiers
1..726
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4799823"

```

/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT      181 a      159 c      226 g      160 t
ORIGIN

alignment_scores:
  Quality: 672.00      Length: 127
  Ratio: 5.291      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x BG714974

Align seg 1/1 to: BG714974 from: 1 to: 726

1 MetGluSerLySGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
219 ATGGATGTCACAGAGAGAGCGGTTAAACAATCTCATCGTGGAAATGT 268
17 lAsnGlnGluAsnAspGluLySGluGluValAlaAsnLySG 34
269 CAACAGGAAATGATGAAAGAGATGAAAGAGGAGCAAGTTGCTAATAAG 318
34 lYGLuProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
319 GGGAGCCCTTGCCCTTACCTTGAATGTAGTGAATCTGTGCTGCTAGA 368
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTr 67
369 GGAACCGTAGGCGGTTCGCGCTTAGCGAGCCCATCTCGCAGTATAGTG 418
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
419 GGACATATGTCATAGCTTGGAGACCCACAGGCAAGGATGAGAGAGAGA 468
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLySGluArg 100
469 ATATGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAG 518
101 GluLySGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
519 GAAACAGTTGATCATAGTCTGCGGCGAGTCAGCAGCTGATCCCTCA 568
117 SHISAspHisAspGluPheCysLeuMet 127
569 CCATGACCATCAGTATGATTTTGCCTTATG 599

seq_name: gb_est2:BF967675

seq documentation_block:
LOCUS      BF967675      793 bp      mRNA      EST      23-JAN-2001
DEFINITION 60228738371 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374549 3',
            mRNA sequence.
ACCESSION  BF967675
VERSION    BF967675.2 GI:12388141
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 793)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/

```

TITLE JOURNAL COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
On Jan 16, 2001 this sequence version replaced gi:12334890.
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM10038 row: f column: 22
High quality sequence start: 20
High quality sequence stop: 762.

FEATURES source

1..793
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4374549"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 185 a 235 c 170 g 203 t
ORIGIN

alignment_scores:
 Quality: 672.00 Length: 127
 Ratio: 5.291 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x BF967675/rev

Align seg 1/1 to reverse of: BF967675 from: 1 to: 793

1 MetGluSerLySGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
|||||
577 ATGGATGTCACAGAGAGAGCGGTTAAACAATCTCATCGTGGAAATGT 528
17 lAsnGlnGluAsnAspGluLySGluGluValAlaAsnLySG 34
|||||
527 CAACAGGAAATGATGAAAGAGATGAAAGAGGAGCAAGTTGCTAATAAG 478
34 lYGLuProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
477 GGGAGCCCTTGCCCTTACCTTGAATGTAGTGAATCTGTGCTGCTAGA 428
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTr 67
|||||
427 GGAACCGTAGGCGGTTCGCGCTTAGCGAGCCCATCTCGCAGTATAGATG 378
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
|||||
377 GGACATATGTCATAGCTTGGAGACCCACAGGCAAGGATGAGAGAGAGA 328
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLySGluArg 100
|||||
327 ATATGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAG 278
101 GluLySGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||

```
277 GAAAGCACTTGCATAGTCTCGGGCAGTCAGCACTGATCCCTCA 228
117 shIsAspHisHisAspGluPheCysLeuMet 127
|||||
227 CCATGACCATCAGATGAGTTTTCCTTATG 197

seq_name: gb_est2:BG707734

seq_documentation_block:
LOCUS BG707734 817 bp mRNA EST 07-MAY-2001
DEFINITION 602671235F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794061 5',
mRNA sequence.
ACCESSION BG707734
VERSION BG707734.1 GI:13984377
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10674 row: n column: 14
High quality sequence stop: 803.
FEATURES
Location/Qualifiers
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4794061"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to R01 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 213 a 171 c 241 g 191 t
ORIGIN
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4794061"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to R01 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
alignment_scores:
Quality: 672.00 Length: 128
Ratio: 5.250 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.219
alignment_block:
US-09-327-750d-31 x BG707734
Align seg 1/1 to: BG707734 from: 1 to: 817
1 MetGluSerLySGluGluAlaLeuAsnValSerGluTyrCysValProArg 17
|||||
221 ATGAGTCCAAAGAGAGACGCGTAAACAATCTCATCGTGGAAATGT 270
|||||
17 lAsnGlnGluAsnAspGluLySGluGluAlaLeuAsnValSerGluTyr 34
|||||
271 CAACACGAGAAATGATGAAAAAGATGAAAGAGGACGAGTTCCTAATAAG 320
|||||
```

```

alignment_scores:
  Quality: 668.00      Length: 128
  Ratio: 5.260         Gaps: 0
  Percent Similarity: 99.219  Percent Identity: 99.219

alignment_block:
US-09-327-750D-31 x BG705843
Align seg 1/1 to: BG705843 from: 1 to: 813:

1 MetGluSerLysGluGluAlaLeuAsnLeuLeuValGluAsnVa 17
218 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCTGGAATGT 267
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
268 CAACAGGAAATGATGATGAAAGATGAAAGGAGCAAGTTGCTAATAAAG 317
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
318 GAGAGCCCTTGCCCTTACCTTGAATGTTAGTGAATCTGTGCTAGTAG 367
51 GlyAsnArgArgArgPheArgValArgGlnProLeuLeuGlnTyrArgTr 67
368 GGAACCCGTAGCGGTTCCCGCTTAGCGAGCCCATCTGCAGTATAGATG 417
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGluA 84
418 GGACATAATGATAGCTTGGAGAGCCAGGCAAGGATGAGAGGAGAGA 467
84 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
468 ATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAG 517
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
518 GAAMGAGCTGAGTCATAGTCTCGGCGAGTCAGCACTGATCCCCCTCA 567
117 sHisAspHisAspGluPheCysLeuMetPro 128
568 CCATGACCATCAGATGATGTTGCTTATGCC 601

seq_name: gb_est1:A1291270

seq_documentation_block:
LOCUS A1291270 511 bp mRNA EST 29-JAN-1999
DEFINITION qm16e06.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1882018 3',
similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD
PROTEIN HGR74, mRNA sequence.
ACCESSION A1291270
VERSION A1291270.1 GI:3934044
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

Insert Length: 742 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 475.

FEATURES

source

Location/Qualifiers
 1..511
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1882018"
 /clone_lib="NCI_CGAP_Lu5"
 /tissue_type="carcinoid"
 /lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 152 a 103 c 155 g 100 t 1 others
 ORIGIN

alignment_scores:

Quality: 666.00 Length: 128
 Ratio: 5.286 Gaps: 0
 Percent Similarity: 98.438 Percent Identity: 98.438

alignment_block:

US-09-327-750D-31 x A1291270

Align seg 1/1 to: A1291270 from: 1 to: 511

1 MetGluSerLysGluGluAlaLeuAsnLeuLeuValGluAsnVa 17
 83 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCTGGAATGT 132
 17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
 133 CAACAGGAAATGATGATGAAAGATGAAAGGAGCAAGTTGCTAATAAAG 182
 34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 183 GAGAGCCCTTGCCCTTACCTTGAATGTTAGTGAATCTGTGCTAGTAG 232
 51 GlyAsnArgArgArgPheArgValArgGlnProLeuLeuGlnTyrArgTr 67
 233 GGAACCCGTAGCGGTTCCCGCTTAGCGAGCCCATCTGCAGTATAGATG 282
 67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGluA 84
 283 GGACATAATGATGATGATGGAAGAGCCAGGCAAGGATGAGAGGAGAGA 332
 84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
 333 ATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAG 382
 101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
 383 GAAACCCAGTCAATAGTCTCGGCGAGTCAGCACTGATCCCCCTCA 432
 117 sHisAspHisAspGluPheCysLeuMetPro 128
 433 CCATGACCATCAGATGATGTTTGCCTTATGCC 466

seq_name: gb_est1:A1291126

seq_documentation_block:

LOCUS A1291126 669 bp mRNA EST 29-JAN-1999
 DEFINITION qm15f02.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1881915 3',
 similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
 (HUMAN); mRNA sequence.
 ACCESSION A1291126

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
333 ATATGGAAGAGTTTGGGAGGAGGTGAGACAGCTGATGGAAAGCTGAGG 382

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
383 GAAAGACAGTGTAGTCATAGTCTGGGCGCAGTCAGCACTGATCCCTCA 432

117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
433 CCATGACCATCAGCATGAGTTTGCCTTATGCC 466

seq_name: gb_est2:BF237433

seq_documentation_block:
LOCUS BF237433 711 bp mRNA EST 14-NOV-2000
DEFINITION 601842108F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4079856 5',
mRNA sequence.
ACCESSION BF237433
VERSION BF237433.1 GI:11151351
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 711)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM938 row: p column: 01
High quality sequence stop: 708.

FEATURES
Location/Qualifiers
1..711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4079856"
/clone_lib="NIH_MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT 184 a 152 c 218 g 157 t
ORIGIN

alignment_scores:
Quality: 664.00 Length: 128
Ratio: 5.270 Gaps: 0
Percent Similarity: 98.438 Percent Identity: 98.438

alignment_block:
US-09-327-750D-31 x A1291126 ..
Align seg 1/1 to: A1291126 from: 1 to: 669

1 MetGluSerLysGluGluArgAlaLeuAsnLeuIleValGluAsnVa 17
83 ATGGAGTCCAAAGAGACGCGTTAAACATCTCATCGGAAATGT 132

17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAsnLysG 34
133 CAACAGGAAATATGATAAAGATGAAAGGACGAGTTCCTAATAAG 182

34 lYcLupProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
183 GGAGGCCCTTGGCCCTTACCTTTGAATGTAGTGAATCTGTGTCCTAGA 232

51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
233 GGAACCGTAGCGGTTCGCGCTTAGGCAGCCCATCTGTCAGTAGATG 282

67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
283 GGACATAATGATAGCTTGGAGAGCCACGACGAGGATGAGAGGAGA 332

seq_name: gb_est2:BF237433

seq_documentation_block:
LOCUS BF237433 711 bp mRNA EST 14-NOV-2000
DEFINITION 601842108F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4079856 5',
mRNA sequence.
ACCESSION BF237433
VERSION BF237433.1 GI:11151351
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 711)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM938 row: p column: 01
High quality sequence stop: 708.

FEATURES
Location/Qualifiers
1..669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1881915"
/clone_lib="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung carcinoid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pTT73 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 186 a 132 c 185 g 164 t 2 others
ORIGIN

alignment_scores:
Quality: 664.00 Length: 128
Ratio: 5.270 Gaps: 0
Percent Similarity: 98.438 Percent Identity: 98.438

alignment_block:
US-09-327-750D-31 x A1291126 ..
Align seg 1/1 to: A1291126 from: 1 to: 669

1 MetGluSerLysGluGluArgAlaLeuAsnLeuIleValGluAsnVa 17
83 ATGGAGTCCAAAGAGACGCGTTAAACATCTCATCGGAAATGT 132

17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAsnLysG 34
133 CAACAGGAAATATGATAAAGATGAAAGGACGAGTTCCTAATAAG 182

34 lYcLupProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
183 GGAGGCCCTTGGCCCTTACCTTTGAATGTAGTGAATCTGTGTCCTAGA 232

51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
233 GGAACCGTAGCGGTTCGCGCTTAGGCAGCCCATCTGTCAGTAGATG 282

67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
283 GGACATAATGATAGCTTGGAGAGCCACGACGAGGATGAGAGGAGA 332

```

|||||
212 ATGGAGTCAAGAGGAAGACGCGCTTAACAATCTCATCGTGGAAATGT 261
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
262 CAACCAAGAAATGATGAAGAAGATGAAGAAGGAGCAAGTGTCTAATAAAG 311
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
312 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCCTAGA 361
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
362 GGAACCCGTAGCGGTTCCGGTTAGCAGCCCATCTGCAGTATAGATG 411
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
|||||
412 GGACATAATGATAGGCTTGAGAGCCACAGGCAAGGATGAGAGAGGAGA 461
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
462 ATATGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 511
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
512 GAAAGCAGTTGAGTCATAGTTTCGGGGCAGTCAGCACTGATCCCCCTCA 561
117 shIsAspHisAspGluPheCysLeuMetPro 128
|||||
562 CCATGACCATCAGCATGAGTTGCTTATGCCCT 595

```

seq_name: gb_est1:BE266012

```

seq_documentation_block: 610 bp mRNA EST 13-JUL-2000
LOCUS BE266012
DEFINITION 601194731Fl NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3538735 5',
mRNA sequence.
ACCESSION BE266012
VERSION BE266012.1 GI:9139583
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 610)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Plate: LCM224 row: m column: 08
High quality sequence stop: 607.

```

FEATURES

```

source
1..610
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:3538735"
  /clone_lib="NIH_MGC_7"
  /tissue_type="small cell carcinoma"
  /cell_line="MGC3"
  /lab_host="DH10B (phage-resistant)"
  /note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site: 2:
  EcoRI; cDNA made by oligo-dT priming. Directionally
  cloned into EcoRI/XhoI sites using the following 5',
  adaptor: GGCAGGAG(G). Size-selected >500bp for average
  insert size 1.8kb. Library constructed by Ling Hong in
  the laboratory of Gerald M. Rubin (University of
  California, Berkeley) using 2AP-cDNA synthesis kit
  (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT
ORIGIN

```

alignment_scores:
  Quality: 663.00      Length: 129
  Ratio: 5.180        Gaps: 1
  Percent Similarity: 99.225  Percent Identity: 98.450

alignment_block:
US-09-327-750D-31 x BE266012
Align seg 1/1 to: BE266012 from: 1 to: 610

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
172 ATGGAGTCAAGAGGAAGACGCGCTTAACAATCTCATCGTGGAAATGT 221
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
222 CAACCAAGAAATGATGAAGAAGATGAAGAAGGAGCAAGTGTCTAATAAAG 271
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
272 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCCTAGA 321
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
322 GGAACCCGTAGCGGTTCCGGTTAGCAGCCCATCTGCAGTATAGATG 371
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
|||||
372 GGACATAATGATAGGCTTGAGAGCCACAGGCAAGGATGAGAGAGGAGA 421
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuAr 100
|||||
422 ATATGCAACAGGATTTGGGAGGAGGTGAGACAGCTGATGCAAAAGCTGAG 471
100 gGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
472 GGAAGAGCAGTTGAGTCATAGTCTCGGGGAGCAGTCAGCACTGATCCCCCTC 521
117 shIsAspHisAspGluPheCysLeuMetPro 128
|||||
522 ACCATGACCATCAGCATGAGTTTGCCTTATGCCCT 556

```


US-08-470-566B-18

alignment_scores:
Quality: 87.00 Length: 126
Ratio: 1.338 Gaps: 6
Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-08-470-566B-18 ..

Align seg 1/1 to: US-08-470-566B-18 from: 1 to: 2004

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
270 GAACGTGAGCATGGAGAGGTGATCTCGAGCCCAACGAGCACTGAGCA 319
31 aasnlysglygluProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
320 ACAGCGTGAGAGCCACTCGAGCACTGAGCACTGAGCACTGAGCACTGAG 365
48 alProArgGlyAsnArgArgPheArgValArgGlnProLeuGln 64
366CGAGGGCCGAGCGTGGAGCGGCATCGTCCCAAGGCATCAG 410
65 TyrArgTrrpAsp..... 68
411 CTTCGGCGTGAGCGTGAACCTACCAGCAGCAGCAGCGTGGCCCGAGGAGT 460
69lleMethHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
461 GGGGACACGACCGCGGACACCGCAGGTCACACCGCCGAGGC.... 506
83 luasnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
507CGGTACCTGAAGCCCAACCTGCGCTACACCAACAGTGGG 545
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
546 CACCGCGCCATCTAGCAGGTGAAGCCCAACCGCAGCAGTTCGTGCT..... 590
112 rThrAspProHisHisAspHisHis 121
591GAACAACGACACCATCGCCACCAT 614

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-18

seq_documentation_block:
Sequence 18, Application US/08469334
Patent No. 5990383

GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5990383el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,334
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 08/463,483
FILING DATE:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2004 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2004
OTHER INFORMATION: /note= "Maize optimized DNA
OTHER INFORMATION: sequence for VIPA(a) 80 kd protein from AB78"

alignment_scores:
Quality: 87.00 Length: 126
Ratio: 1.338 Gaps: 6
Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-08-469-334-18 ..

Align seg 1/1 to: US-08-469-334-18 from: 1 to: 2004
15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
270 GAACGTGAGCATGGAGAGGTGATCTCGAGCCCAACGAGCACTGAGCA 319
31 aasnlysglygluProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
320 ACAGCGTGAGAGCCACTCGAGCACTGAGCACTGAGCACTGAGCACTGAG 365
48 alProArgGlyAsnArgArgPheArgValArgGlnProLeuGln 64
366CGAGGGCCGAGCGTGGAGCGGCATCGTCCCAAGGCATCAG 410
65 TyrArgTrrpAsp..... 68
411 CTTCGGCGTGAGCGTGAACCTACCAGCAGCAGCAGCGTGGCCCGAGGAGT 460
69lleMethHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
461 GGGGACACGACCGCGGACACCGCAGGTCACACCGCCGAGGC.... 506
83 luasnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
507CGGTACCTGAAGCCCAACCGCAGCAGTTCGTGCT..... 545

```
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValise 112
      ::::::::::::::
      ::::::::::::::
546 CACCGCGGCCATCTACGACGTGAAGCCACCCACCGCTTCGTGCT..... 590
112 rThrAspProHisHisAspHis 121
      ::::::::::::::
591 ....GACACAGCACCATCGCCACCAT 614

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-300-529-18
seq_documentation_block:
; Sequence 18, Application US/09300529
; Patent No. 6066783
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: Genes Encoding Insecticidal Proteins
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066783artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300.529
; FILING DATE: TBA
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/469,334
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-19506L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2004
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP1A(a) 80 kd protein from AB78"
; US-09-300-529-18

alignment_scores:
  Quality: 87.00 Length: 126
  Ratio: 1.338 Gaps: 6
Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:
US-09-327-750d-31 x US-09-300-529-18

Align seg 1/1 to: US-09-300-529-18 from: 1 to: 2004

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGlnValAl 31
   |||:::  ::::::::::::::
270 GAACGTGACATGGAGAGGTGATCTTGAGCCCAAGGAGAACTGAGCA 319
   |||:::  ::::::::::::::
31 AsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysV 48
   :::::  |||:::  ::::::::::::::
320 ACAGCGTGGAGAGCCACTCGAGGACCACTGAGCTACACCAACAC... 365
   :::::  |||:::  ::::::::::::::
48 alProArgGlyAsnArgArgPheArgValArgGlnProLeuLeuGln 64
   |||:::  ::::::::::::::
366 .....CGAGGGCGCCAGCGTGGAGCGCGCATCGTCCCAAGGGCATCAG 410
   |||:::  ::::::::::::::
65 TyrArgTirAsp..... 68
   |||  :::
411 CTTGCGCGTGAAGCTGAAGTACCAGCAGCAGGAGCGTGGCCAGGAGT 460
   ::::::::::::::
69 .....IleMethisArgLeuGlyGluPro.....GlnAlaArgMetArgLuc 83
   ::::::::::::::
461 GGGGACCAACCGCGGCAACACCGCCGTTCAACACCGCCAGCGC... 506
   |||:::  |||:::  ::::::::::::::
83 LuAsnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
   |||:::  |||:::  ::::::::::::::
507 .....CGGCTACCTGAGCGCCACGCTGCGCTACCAACACGTTGG 545
   ::::::::::::::
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValise 112
   ::::::::::::::
546 CACCGCGGCCATCTACGACGTGAAGCCACCCACCGCTTCGTGCT..... 590
   ::::::::::::::
112 rThrAspProHisHisAspHisHis 121
   ::::::::::::::
591 ....GACACAGCACCATCGCCACCAT 614
   ::::::::::::::

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-35
seq_documentation_block:
; Sequence 35, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
```


ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,033
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: P-40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLv3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 2576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 9...2564
OTHER INFORMATION: /note= "Maize optimized sequence
OTHER INFORMATION: encoding VIPA(a) with the Bacillus secretion signal removed
OTHER INFORMATION: contained in pCIB5526"
US-08-471-033-35

alignment_scores:
Quality: 87.00 Length: 126
Ratio: 1.338 Gaps: 6
Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-08-471-033-35 ..
Align seg 1/1 to: US-08-471-033-35 from: 1 to: 2576

15 GluAsnValAsnGlnGluAsnAspGluLysGluGlnValAl 31
833 GAACGTGAGCAGGAGGAGGTCTCTGAGCCCAACGAGCACTGAGCA 882
31 aAsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
883 ACAGCTGGAGAGCCACTCGAGCAACCACTGGAGCTACACCAACAC... 928
48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
929CGAGGCGCCAGCGTGGAGCGCGCATCGGTCCCAAGGCGCATCAG 973
65 TyrArgTrpAsp..... 68
974 CTTCCGGCTGAGGTGAACCTACACAGCAGCAGCAGCGTGGCCAGGAGT 1023
69IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
1024 GGGGACACGACCGGCAACACACAGCAGCAGTTCACACACCGCGCGC... 1069

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
1070CGGTACTACCAACCGCAACGTGCGCTACAAACAGTGGG 1108
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
1109 CACGGGCGCATCTACGACGTGAAGCCACACACAGCTTCGTGCT..... 1153
112 rThrAspProHisHisHisHisHis 121
1154GAACACGACACCATCGCCACCAT 1177
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-044-35
seq_documentation_block:
Sequence 35, Application US/08471044
Patent No. 5840868
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Kozziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,044
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/463,483
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 23-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLv3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 2576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"


```

; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2655 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2652
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for 100 kd VIP1(a) protein from AB78"
US-08-471-033-17

alignment_scores:
    Quality:      87.00      Length:      126
    Ratio:        1.338      Gaps:          6
    Percent Similarity: 51.587   Percent Identity: 27.778

alignment_block:
US-09-327-750D-31.x US-08-471-033-17 ..

Align seg 1/1 to: US-08-471-033-17 from: 1 to: 2655

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
||||| : : : ||||| : : : : : : : : |
921 GRACGTGAGCTGGAGAGGTCCTCGAGCCCAAGGAACTTGACGA 970

31 aaSnLysGlyGluProLeuAlaLeuProLeuAsnValSerGlyTyrCysV 48
: : : : : ||||| : : : : : : : : :
971 ACAGCGTTGAGAGCCACTCGAGCACCACTGGAGCTACACCACAC.... 1016

48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
||||| : : : : : ||||| : : : : :
1017 .....CGAGGGGCCAGCGTGGAGCCGGCATCGTCCCAGGCGATCAG 1061

65 TyrArgTrpAsp..... 68
||| : : :

1062 CTTCCGGCTGAGCGTAACCTACCAGCACAGCGAGCGTGGCCCCAGGAGT 1111

69 ....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
: : : ||||| : : : : : ||||| : : : : :
1112 GGGCCACCAGCCGGACACACCAGCCAGTTCAACACCGCCAGCG.... 1157

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
||||| : : : : : ||||| : : : : :
1158 .....CGGTACCTGAACGCCAACCGTCGCCTACAAACAACGTGGG 1196

100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
: : : : : ||||| : : : : :
1197 CACGGGGCCNCTACGAGCTGAAGCCACCACCAGCTTCGTCT..... 1241

112 rThrAspProProHisHisAspHisHis 121
: : : ||||| : : : : :
1242 ....GAACACGACACCATTCGCCCAT 1265

```

OM of: US-09-327-750d-31 to: N_Geneseq_1101.* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL-frame-p2n.model -DEV-xlh
-O/cgn2_1/USPTO.spool/US09327750/runat_11032002.101154_20362/app_query.fasta_1.1472
-DB-N_Geneseq_1101 -OFMT-fastap -SUFFIX-rng -GAPOP-12.000
-GAPEXT-4.000 -MINMATCH-0.100 -LOOPCL-0.000 -LOOPEXT-0.000
-GAPOP-4.500 -GAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500
-GAPOP-6.000 -GAPEXT-7.000 -YGAPOP-10.000 -YGAPEXT-0.500
-DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-blossum62
-TRANS-human40 cdi -LIST-45 -DOCALIGN-200 -THR_SCORE-pct
-THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pfs
-NORM-ext -HEAPSIZ-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09327750_cgn1_1_330 -NCPU-6 -ICPU-3 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIME-30 -NO_XLPXY -WAIT -THREADS-1

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810

seq_documentation_block:
ID AAH75810 standard; cDNA; 767 BP.
XX
AC AAH75810;

seq_documentation_block:
ID AAH75810 standard; cDNA; 767 BP.
XX
AC AAH75810;

XX 17-OCT-2001 (first entry)
XX Human X chromosome linked gene expression protein 14 coding sequence.

XX 17-OCT-2001 (first entry)
XX Human X chromosome linked gene expression protein 14 coding sequence.

DE Human; X chromosome linked gene expression protein 14; cancer;
KW HIV infection; cytostatic; anti-HIV; chromosome X; ss.

DE Human; X chromosome linked gene expression protein 14; cancer;
KW HIV infection; cytostatic; anti-HIV; chromosome X; ss.

XX Homo sapiens.
XX CN1296969-A.

XX Homo sapiens.
XX CN1296969-A.

XX 30-MAY-2001.
XX 23-NOV-1999; 99CN-0124078.

XX 30-MAY-2001.
XX 23-NOV-1999; 99CN-0124078.

XX 23-NOV-1999; 99CN-0124078.
XX (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX 23-NOV-1999; 99CN-0124078.
XX (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX Mao Y, Xie Y;
XX WPI; 2001-483897/53.

XX Mao Y, Xie Y;
XX WPI; 2001-483897/53.

DR P-PSDB; AAG66407.
PT Polypeptide-human X chromosome linked gene expression protein 14 and
PT polynucleotide for coding said polypeptide -

DR P-PSDB; AAG66407.
PT Polypeptide-human X chromosome linked gene expression protein 14 and
PT polynucleotide for coding said polypeptide -

XX Claim 6; Page 24 (Disclosure); 31pp; Chinese.
XX The present sequence is the coding sequence for human X chromosome linked
XX gene expression protein 14. The protein and coding sequence are useful
XX for treating diseases e.g. cancer and HIV infection.

XX Claim 6; Page 24 (Disclosure); 31pp; Chinese.
XX The present sequence is the coding sequence for human X chromosome linked
XX gene expression protein 14. The protein and coding sequence are useful
XX for treating diseases e.g. cancer and HIV infection.

XX Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;
SQ

XX Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;
SQ

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x AAH75810
Align seg 1/1 to: AAH75810 from: 1 to: 767

alignment_block:
US-09-327-750D-31 x AAH75810
Align seg 1/1 to: AAH75810 from: 1 to: 767

1 MetGluSerLysGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
101 ATGGAGTCCAAAGAGGACGAGCGCTTAAACAATCTCATGCTGGAATGT 150

1 MetGluSerLysGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
101 ATGGAGTCCAAAGAGGACGAGCGCTTAAACAATCTCATGCTGGAATGT 150

17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
151 CAACAGAGAAATGATGAAAGATGAAAGAGGAGGAGGAGGAGGAGGAGG 200

17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
151 CAACAGAGAAATGATGAAAGATGAAAGAGGAGGAGGAGGAGGAGGAGG 200

34 IyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
201 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGTGCTAGA 250

34 IyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
201 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGTGCTAGA 250

51 GlyAsnArgArgArgPheArgValArgGlnProIleuGluTyrArgTr 67
|||||

51 GlyAsnArgArgArgPheArgValArgGlnProIleuGluTyrArgTr 67
|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

251 GGAACCGTAGGGGTTCCGCTTAGCAGCCCACTCTCGAGTATAGATG 300
 67 PASPIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
 301 GGACATAATGCATAGCTTGGAGAGCCACAGGCAAGCATGAGAGAGAGA 350
 84 sMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
 351 ATATGGAAGAGATTGGGAGAGGTGAGACAGCTGATGGAAGAGCTGAGG 400
 101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
 401 GAAAAGCAGCTTCAGTCATAGTCTCGGCGCAGTCAGCACTATCCCTCA 450
 117 SHISAspHisHisAspGluPheCysLeuMetPro 128
 451 CCATGACCATCATCAGTATGTTTGCCTTATGCCCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI58581

seq_documentation_block:

ID AAI58581 standard; cDNA; 862 BP.

AC AAI58581;

DF 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 784.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.

DR P-PSDB; AAM39425.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Claim 1: SEQ ID NO 784; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

SQ Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores

Quality: 679.00 Length: 128
 Ratio: 5.305 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x AAI58581

Align seg 1/1 to: AAI58581 from: 1 to: 862

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
 226 ATGGAGTCCARAGAGGAGGAGCGTTAAACAACTCTCATCTGGGAAAATGT 275
 17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
 276 CAACGAGGAAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTAAATAAG 325
 34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 326 GGGAGCCCTTGCCCTTACCTTGAATGTTAGTGAATCTGTGTCCTAGA 375
 51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
 376 GGAACCCGTAGGCGGTTCCGGTTAGGAGCCCATCTCGCAGTATAGATG 425
 67 PASPIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
 426 GCACATATGATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 475
 84 sMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
 476 ATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 525
 101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
 526 GAAAAGCAGTTGAGTCATAGTTTGGGCGCAGTCAGCACTATCCCTCA 575
 117 SHISAspHisHisAspGluPheCysLeuMetPro 128
 576 CCATGACCATCATCAGTATGTTTGCCTTATGCCCC 609

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611

seq_documentation_block:

ID AAF59611 standard; cDNA; 898 BP.

AC AAF59611;

DT 24-APR-2001 (first entry)

DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.

KW Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;

menstrual cycle disorder; bacterial infection; ss.

Homo sapiens.

WO200107471-A2.

01-FEB-2001.

21-JUL-2000; 2000WO-US19948.

21-JUL-1999; 99US-0145075.

08-SEP-1999; 99US-0153129.

10-NOV-1999; 99US-0164647.

(INCY-) INCYTE GENOMICS INC.

Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman D;
Azimzai Y, Yang P, Lu DAM, Baughn MR, Patterson C, Shah P;
WPI: 2001-112727/12.
P-PSDB: AAB60474.

Human cell cycle and proliferation proteins and polynucleotides are used to treat, diagnose and prevent immune, developmental and cell signaling disorders and cell proliferative disorders including cancer -

Claim 5; Page 181-182; 205pp; English.

Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506. CCYPR and agonists of CCYPR are used to treat diseases or conditions associated with decreased expression of functional CCYPR, while CCYPR antagonists are used to treat diseases or conditions associated with overexpression of functional CCYPR. Monoclonal or polyclonal antibodies to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect compounds e.g., antibodies, oligonucleotides and proteins (receptors) that specifically bind to CCYPR, and in drug screening methods to identify compounds that modulate the activity of CCYPR. CCYPR nucleotides can be used to generate transgenic animal models of human disease, and can be used in gene therapy in target cells with genetic abnormalities with respect to the expression of CCYPR for the treatment or prevention of a disorder associated with CCYPR. Diseases which can be diagnosed, treated and prevented using CCYPR proteins, nucleic acids, agonists or antagonists include immune, developmental and cell signalling disorders, and cell proliferative disorders including cancer. Specific examples of these disorders include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies, diabetes mellitus, disorders of the menstrual cycle and infections caused by bacteria.

Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x AAF59611 ..

Align seg 1/1 to: AAF59611 from: 1 to: 898

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuValGluAsnVa 17
|||||
254 ATGGAGTCCAAAGAGACGAGCGTAAACAATCTCATCGTGGAAATGT 303
17 laaGlnGluAsnAspGluLysAspGluLysGluValAlaAlaAsnLysG 34
|||||
304 CAACACAGAAATATGTAAGAAAGATGAAGAGACGACGAGTCTAATAAG 353
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

|||||
354 GGAGCCCTGGCCCTACCTTGAATGTAGTGAATCTGTGTGCTCTAGA 403
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
404 GGAACCGTAGCGGTTCGGCTTAGGCAGCCCATCTCTGCAGTAGATAG 453
67 pasPileMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 84
|||||
454 GGACATATTCATAGCTTGGAGACCCACAGCAGGATGAGAGAGAGAGA 503
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
504 ATATGGAAGAGATTGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAG 553
101 GluLysGluLeuSerHisSerHisLeuArgAlaValSerThrAspProH 117
554 GAAACGAGTTGAGTCATAGTCTGCGGCGAGTCAGCACTGATCCCTCA 603
117 shiAspHisAspGluPheCysLeuMetPro 128
|||||
604 CCATGACCATCAGGATGAGTTTTCCTTATGCC 637

seq_name: /SIDS2/cgdata/geneseq/geneseq/NA2001.DAT:AAI60367

seq_documentation_block:
ID AAI60367 standard; cDNA; 858 BP.
XX
AC AAI60367;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4356.
XX
KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
WPI: 2001-442253/47.
DR P-PSDB; AAM41211.
XX
Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PT
PS Claim 1; SEQ ID NO 4356; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AA38642-AA42213), with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

SQ Sequence 858 BP; 236 A; 175 C; 241 G; 206 T; 0 other;

alignment_scores:
 Quality: 666.00 Length: 129
 Ratio: 5.203 Gaps: 1
 Percent Similarity: 99.225 Percent Identity: 99.225

alignment_block:

US-09-327-750D-31 x AA160367

Align seg 1/1 to: AA160367 from: 1 to: 858

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
 |||||
 215 ATGGAGTCCAAAGAGGACGCGTTAAACAATCTCATCTGCGAAATGT 264
 17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
 |||||
 265 CAACACGAGAAATGATCAAAAGATGAAAGGAGCAAGTTGCTATAAAG 314
 34 IyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 |||||
 315 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCTAGA 364
 51 GlyAsnArgArgPheArgValArgGlnProIleLeuGluTyrArgTr 67
 |||||
 365 GGAACCCGTAGCGGTTCCGCTTAGCGAGCCCATCTCGAGTATAGTG 414
 67 PaspIleMethHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
 |||||
 415 GGACATATGATAGCTAGGCTTGAGAGCCACAGGAGGATGAGAGAGAGA 464
 84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
 |||||
 465 ATATGGAAGGATTGGGAGGAGTGAGACAGCTGATGGAAGAGCTGAGG 514
 101 GlnLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH 117
 |||||
 515 GAAAGAGCTTGATGATAGTCTGCGGGCAGTCAGCAGCTGATGATCCCTCA 564
 117 SHISAspHisAspGluPheCysLeuMetPro 128
 |||||
 565 CCATGACCATCAGCATGAGTTTGGCCCTTATGCC 599

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC03880

seq_documentation_block:

ID AAC03880 standard; cDNA: 662 BP.

XX AAC03880;

XX AC

XX XX

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 3878.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX KW

KW

XX

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PA

XX

PI

XX

DR

XX

DR

XX

PT

XX

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

gene therapy; chromosome mapping; ss.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI; 2000-500381/45.

P-PSDB; AAG03874.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for

obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

diagnostic, forensic, gene therapy and chromosome mapping procedures -

Claim 1; SEQ ID 3878; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from

mRNAs encoding secreted proteins. An ORF has been identified within the

sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

derived from 30 different tissues. EST sequences usually correspond

mainly to the 3' untranslated region (UTR) of the mRNA because they are

often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

well suited for isolating cDNA sequences derived from the 5' ends of

mRNAs and even in those cases where longer cDNA sequences have been

obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

mRNAs with intact 5' ends and can therefore be used to obtain full length

cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

gene therapy and chromosome mapping procedures. They are used to obtain

upstream regulatory sequences and to design expression and secretion

vectors.

Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

alignment_scores:

Quality: 649.00 Length: 128

Ratio: 5.234 Gaps: 0

Percent Similarity: 96.875 Percent Identity: 95.312

alignment_block:

US-09-327-750D-31 x AAC03880

Align seg 1/1 to: AAC03880 from: 1 to: 662

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

|||||

207 ATGGAGTCCAAAGAGGACGCGTTAAACAATCTCATCTGCGAAATGT 256

17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

|||||

257 CAACACGAGAAATGATCAAAAGATGAAAGGAGCAAGTTGCTATAAAG 306

34 IyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

|||||

307 GGGAGCCCTTGGCCCTACCTTTGATGTTGTTGAATACTGTGTGCTAGA 356

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGluTyrArgTr 67

|||||

357 GGAATCGTAGCGGTTCCGCGTTAGCAGCCCATCTCTGCAGTATAGTG 406

67 PaspIleMethHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84

|||||

407 GGATATGATGATAGCTAGGCTTGAGAGAACACAGGAGGATGAGAGAGA 456

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
 |||||
 457 ATATGGAAGAGTGGGAGGAGGTGAGACAGCTGATGGAAAGCTGAGG 506
 |||||
 101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
 |||||
 507 GAAAGCAGCTTGGTCACTAGTCTGGGGCAGCTCAGCAGTACCCCTCA 556
 |||||
 117 sHisAspHisAspGluPheCysLeuMetPro 128
 |||||
 557 CCATGACCATCATGATGATGATTTGCNNWATGCC 590
 |||||

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC03879

seq_documentation_block:

ID AAC03879 standard; cDNA; 698 BP.

AC AAC03879;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC03879

seq_documentation_block:

ID AAC03879 standard; cDNA; 698 BP.

AC AAC03879;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -
 Claim 1; SEQ ID 3877; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from
 mRNAs encoding secreted proteins. An ORF has been identified within the
 sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 derived from 30 different tissues. EST sequences usually correspond
 mainly to the 3' untranslated region (UTR) of the mRNA because they are
 often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 well suited for isolating cDNA sequences derived from the 5' ends of
 mRNAs and even in those cases where longer cDNA sequences have been
 obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 mRNAs with intact 5' ends and can therefore be used to obtain full length
 cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 gene therapy and chromosome mapping procedures. They are used to obtain
 upstream regulatory sequences and to design expression and secretion
 vectors.

Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:

Quality: 649.00 Length: 128
 Ratio: 5.234 Caps: 0
 Percent Similarity: 96.875 Percent Identity: 95.312

alignment_block:

US-09-327-750D-31 x AAC03879

Align seg 1/1 to: AAC03879 from: 1 to: 698

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
 |||||
 243 ATGAGTCCAAAGAGGAGGAGGTTAAACAATCTCATCGTGGAATGT 292
 |||||
 17 lasGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
 |||||
 293 CAACCCAGGAAATGATGAAAAAGATGAAAGGAGCAAGTGCCTAATAAG 342
 |||||
 34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 |||||
 343 GGGAGCCCTGGCCCTACCTTTGATGTRGTGTAATACTGTGCTCTAGA 392
 |||||
 51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
 |||||
 393 GGAATCGTAGGCGGTTCCGCTTAGGCGCCATCCCTGCAGTATAGATG 442
 |||||
 67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
 |||||
 443 GGATATGATCCTAGGCTTGGAGAACACAGGAGGATGAGAGAGAGA 492
 |||||
 84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
 |||||
 493 ATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGG 542
 |||||
 101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
 |||||
 543 GAAAGCAGTTGATGATCATAGTCTGGGCGGAGTGCAGCACTGACCCCTCA 592
 |||||
 117 sHisAspHisAspGluPheCysLeuMetPro 128
 |||||
 593 CCATGACCATCATGATGATGATTTGCNNWATGCC 626
 |||||

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC10889

seq_documentation_block:

ID AAC10889 standard; cDNA; 692 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

06-OCT-2000 (first entry)

Human secreted protein 5' EST, SEQ ID NO: 14964.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 gene therapy; chromosome mapping; ss.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI; 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -
 Claim 1; SEQ ID 14964; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from

XX OS XX PN XX PD XX PF XX PR XX PA XX PI XX

PI Mao Y, Xie Y;
XX WPI; 2001-397944/42.
DR P-PSDB: AAB99224.
DR

DR P-PSDB; AAB99224.

XX
PS Claim 5; Page 22; 30pp; Chinese.

The present sequence is the coding sequence for a human brain-expressed x-linked protein (hBx). hBx and its coding sequence are useful in the diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer, tumours, deafness, x-chromosome-binding mental retardation and lissencephalous disease. hBx is also useful for screening mimics, agonists, or inhibitors, and in peptide fingerprinting identification. hBx coding sequence can be used as primers or probes, or in producing gene chips or microarrays.

CC tumours, deafness, X-chromosome-binding mental retardation
CC lissencephalous disease. hhex is also useful for screening
CC agonists, or inhibitors, and in peptide fingerprinting ide
CC hhex coding sequence can be used as primers or probes, or
CC gene chips or microarrays.

```

alignment_scores:
  Quality: 592.50      Length: 128
  Ratio: 4.857         Gaps: 1
  Percent Similarity: 95.312  Percent Identity: 87.500

```

```

alignment_block:
US-09-327-750D-31 x AAH45143
Ratio: 4.957
Percent Similarity: 95.312
Gaps: 1
Percent Identity: 87.500

```

```

US-09-327-750D-31 x AAH45143  ..
Align seq 1/1 to: AAH45143 from: 1 to: 792

```

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
173 ATGGACTCCAAGAGAAACTAGCAGTAAACAGTCTCAGCATGGAAAATGC 222

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
:|||||
::|||
223 CAACCAAGAAAT.....GAAGAAAGGAGCAAGTTCCTAATAAG 263

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||
|||
264 GGGAGCCCTTGCCCCCTCCCCTTTGGATGCTGGTGAATACTGTGTGCCTAGA 313
|||||

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
314 GGAATCGTAGGGCGTTCCGGCTTAGGCAGCCCATCTTCAGTATAGATG 363

67 pAspIleMethHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84

07 pspiremetrisi ypeuuyoyufroginaiaia, gmetaiygiugua 84
 |||||:::|||||
 364 GGATATGATGCATAGGCTTGGAGAACCACAGGCAAGGATGAGAGAAGA 413

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetCiuLysLeuArg 100
|||||

414 ATATGGAAGGATTGGGGGGGGTGGACAGCTGATGGAAGCTGAGG 463

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProHis 117
117
464 GAAACAGCTTGAGTCATAGTCTCGGGGGCAGTCAGCAGTCAGCCCCCTCA 513

Percent Similarity: 91.667 Percent Identity: 84.722

alignment_block:

US-09-327-750D-31 x AAC06180

Align seg 1/1 to: AAC06180 from: 1 to: 457

```

1 MetGluSerLysGluGluAlaLeuAsnAsnLeuValGluAsnVa 17
|||||
243 ATGGAGTCCAAAGAGGACGAGCTTAAACAATCTCATCGTGAATAATGT 292
|||||
17 lasnGlnGluAsnAspGluLysAspGluGluGlnValAlaAsnLysG 34
|||||
293 CACACAGGAATGATGAAAGATGAAAGGAGCAAGTTGCTAATAAAG 342
|||||
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
343 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCTAGA 392
|||||
51 GlyAsnArgArgGluGluAlaLeuGlnProLeuLeuGln.TyrArgT 67
|||||
393 GGAACCCGTAG.CGTTCCGCTTAGGCACACAGTATTAGAGCCCGCT 441
|||||
67 rpAspIleMethis 71
::: |||
442 GCCCAGTGACACAT 455

```

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC06184

seq_documentation_block:

ID AAC06184 standard; cDNA; 451 BP.

AC AAC06184;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 10259.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST.) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 1; SEQ ID 10259; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be

CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. CC They are used to obtain upstream regulatory sequences and to design CC expression and secretion vectors.

XX SQ Sequence 451 BP; 117 A; 110 C; 131 G; 91 T; 2 other;

alignment_scores:

Quality: 261.00 Length: 69
Ratio: 4.143 Gaps: 1
Percent Similarity: 91.304 Percent Identity: 81.159

alignment_block:

US-09-327-750D-31 x AAC06184

Align seg 1/1 to: AAC06184 from: 1 to: 451

```

4 LysGluGluArgAlaLeuAsnAsnLeuValGluAsnValAsnGlnG1 20
|||||
246 CAAAGAGAACGAGCGTTAAACAATCTCATCGTGAATAATGTCAACACGGA 295
|||||
20 uAsnAspGluLysAspGluLysGluGlnValAlaAsnLysGlyGluProL 37
|||||
296 AAATGATGAAAGATGAAAGGAGCAAGTTGCTAATAAAGGGGAGCCCT 345
|||||
37 euAlaLeuProLeuAsnValSerGluTyrCysValProArgGlyAsnArg 53
|||||
346 TGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCTAGAGGAACCGT 395
|||||
54 ArgArgPheArgValArgGlnProLeuLeuGln.TyrArgTtpAspIleM 70
|||||
396 AG.CGTTCCGCTTAGGCACACAGTATTAGAGCCACCGCTGCCCACTGA 444
|||||
70 etHis 71
|||
445 CACAT 449

```

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAP23529

seq_documentation_block:

ID AAP23529 standard; DNA; 891 BP.

AC AAP23529;

DT 22-MAR-2001 (first entry)

DE Human NADE DNA.

KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis; neurogenetic disease; NF-kappaB; ds.

OS Homo sapiens.

PN WO200075278-A2.

PD 14-DEC-2000.

PF 07-JUN-2000; 2000WO-US15621.

PR 07-JUN-1999; 99US-0327750.

PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Sato T;

DR WPI; 2001-061707/07.

PT New p75-neurotrophin receptor-associated cell death executor (NADE) and the gene encoding NADE, useful for modulating the activity of p75NTR and for detecting neurodegenerative diseases.

PS Disclosure; Fig 1; 134pp; English.

332 GAAATATCAGCGGGGCGAGTTAGCGACTTGTCCCTAATTTTCGATG 381
 67 pAspIleMetHisArg...LeuGlyGluProGlnAlaArgMetArgGluG 83
 382 GGCCATACCTATAGGCATATTGACGCAATGAAGCGAGA.....G 422
 83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
 423 ATGATGTAGAAAGGTTTGTAGGCGAGATGATGGAATCAAGAGAAAGACT 472
 100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 473 AGGGAACAGCAGATGAGCATATATGCGCTTCCAAACTCCTGAACCT.. 520
 116 ohHisAspHisHisAspGluPheCysLeuMetPro 128
 521GACAACCATTTATGACTTTTGCCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750

seq_documentation_block:

ID AAH13750 standard; cDNA; 1229 BP.

AAH13750;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:10656.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Human sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX
 SQ Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:

Quality: 187.00 Length: 129

Ratio: 2.226 Gaps: 4

Percent Similarity: 65.116 Percent Identity: 37.984

alignment_block:

US-09-327-750D,31 x AAH13750 ..

Align seg 1/1 to: AAH13750 from: 1 to: 1229

1 MetGluSerLysGluGluArgAlaLeuAsnLeuIleValGluAsnVa 17

191 ATGGAGTCCAAAGAGGAACTAGCGGCAACAACATCTCAACGGGGAATGC 240

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

241 CCAACAAGAAAC.....GAAGGAGGGGAGCAGGCCCCACGCAGA 281

34 lYGLuProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

282 ATGAGAGAATCCCGCCATTTGGGGGGGTGAAGGCCAGAACCTTGA 331

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

332 GGAATATCAGCGGGGCGAGTTAGCGACTTGTCCCTAATTTTCGATG 381

67 pAspIleMetHisArg...LeuGlyGluProGlnAlaArgMetArgGluG 83

382 GGCCATACCTATAGGCATATTGACGCAATGAAGCGAGA.....G 422

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

423 ATGATGTAGAAAGGTTTGTAGGCGAGATGATGGAATCAAGAGAAAGACT 472

100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116

473 AGGGAACAGCAGATGAGCATATATGCGCTTCCAAACTCCTGNACCT.. 520

116 ohHisAspHisHisAspGluPheCysLeuMetPro 128

521GACAACCATTTATGACTTTTGCCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750

seq_documentation_block:

ID AAC85548 standard; cDNA; 1364 BP.

AAH13750;

04-JUN-2001 (first entry)

Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;

cell proliferation; Alzheimer's disease; schizophrenic disorder;

arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.

Human sapiens.

Key Location/Qualifiers

Key 250..612

Key 250..612

441 GGCCATACCTAATAGGCATATTGAGCACAATGAGCGAGA.....G 481
83 luasnMetGluArgGileGlyGluGluValArgGlnLeuMetGluLysLeu 99
482 ATGATGTAGAAAGCTTTTGTAGGCAGATGATGTAATCAAGAGAAGACT 531
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
532 AGGACACAGCAGATGAGGACACTATATGCGCTTCCAACTCTGAACCT.. 579
116 OHHisHisAspHisHisAspGluPheCysLeuMetPro 128
580GACACCATTATGACTTTTGCTCATACCT 609

FT /*tag= a
FT /product= "CDIFF-4"
XX WO200119860-A2.
XX 22-MAR-2001.
XX 14-SEP-2000; 2000WO-US25435.
XX 15-SEP-1999; 99US-0154140.
XX 06-DEC-1999; 99US-0169155.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
XX Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;
XX WPI: 2001-211447/21.
XX P-PSDB: AAB47126.
XX Isolated polypeptides and polynucleotides involved in cell
XX differentiation are used for treatment, prevention and diagnosis of
XX cell proliferative, developmental and neurological disorders e.g.
XX cancer and Alzheimer's disease.
XX Claim 5; Page 121; 137pp; English.
XX The sequences given in AAC85545-72 encode human polypeptides involved
XX in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
XX these are used to treat a disease or condition associated with
XX decreased expression of functional CDIFF. An antagonist of CDIFF is
XX used to treat a disease or condition associated with over expression
XX of functional CDIFF. CDIFF polypeptides may be used for the treatment,
XX prevention and diagnosis of cell proliferative, developmental and
XX neurological disorders, such as Alzheimer's disease, schizophrenia
XX and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
XX REX-3. This sequence maps to chromosome 1 within the interval from
XX 152.2 to 157.4 centimorgans, to chromosome 3 within the interval from
XX 157.4 to 158.0 centimorgans, and to the X chromosome within the interval
XX from 104.9 to 150.3 centimorgans.
XX Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:
Quality: 187.00 Length: 129
Ratio: 2.226 Gaps: 4
Percent Similarity: 65.116 Percent Identity: 37.984
alignment_block:
US-09-327-750D-31 x AAC85548 ..
Align seg 1/1 to: AAC85548 from: 1 to: 1364
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuValGluAsnVa 17
250 ATGGAGTCCAAAGAGTAAGTAGCGCAACAATCTCAACGGGGGAAATGC 299
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
300 CCAACAAGAAAC.....GAAGAGGGGAGCAGGCCGCCAGCAGA 340
34 lYGluproLeuAlaLeuproLeuAsnValSerGluTyrCysValProArg 50
341 ATGAAGAAGAAATCCCGCATTTGGGAGGGGTGAAGGCCAGAGCTGGA 390
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
391 GGAATATACAGCGGGGGGAGTAGCGACTTGCTCCCTAATTTTCGATG 440
67 pAspIleMetHisArg...LeuGlyGluProGlnAlaArgMetArgGluG 83

OM of: US-09-327-750D-31 to: GenEmbl.* out_format : pfs

Date: Mar 11, 2002 3:34 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODELFramet_p2n.model -DEV-xlh
-D/cgcn2_1/USPto.spool/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
-QB/GenEmbl -QFMT-fastap -SUFFIX-rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750 -CGCN1_1.8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-31

Query length: 128

Database: GenEmbl.*

Database sequences: 1472140

Database length: -341344837

Search time (sec): 4557.230000

```
score_list:
Sequence      Strd Orig  2Score  EScore Len  Documentation
gb.pr:AF251053 + 679.00 1230.71 2.8e-60 744  AF251053 Homo sapiens X-linked
gb.pr:AF251053 + 679.00 1229.04 3.5e-60 898  AX078272 Sequence 76 from Paten
gb.pr:HSV870H8 + 679.00 1197.46 2.0e-58 31321  Z70233 Human DNA sequence fr
gb.pr:AL133348 - 679.00 1195.16 2.7e-58 40584  AL133348 Human DNA sequence
gb.hlg:HSU80B1 + 679.00 1195.06 2.7e-58 41029  AL022169 Homo sapiens chromos
gb.pr:AF183416 + 606.50 1099.72 5.6e-53 642  AF183416 Homo sapiens ovarian
gb.pr:AF220189 + 606.50 1097.45 7.5e-53 828  AF220189 Homo sapiens uncharac
gb.pr:HS198P4 - 606.50 1063.99 5.5e-51 35714  AL008708 Human DNA sequence
gb.pr:AF237783 + 592.50 1072.31 1.9e-51 731  AF237783 Homo sapiens brain exp
gb.ro:AF097439 + 452.00 908.13 2.6e-42 785  AF097439 Mus musculus brain exp
gb.ro:AF097437 + 452.00 866.54 1.2e-36 2269  AF097437 Mus musculus Bex1 prc
gb.ro:AF097438 + 450.00 811.78 6.1e-37 835  AF097438 Mus musculus brain exp
gb.ro:AF051347 + 450.00 811.74 6.1e-37 838  AF051347 Mus musculus REX-3 mRN
gb.ro:BC003254 + 388.00 687.56 5.1e-30 2901  BC003254 Mus musculus, Similar
gb.sts:G24641 - 339.00 613.70 6.6e-26 504  G24641 human STS WI-11354, sequ
gb.pr:HSV351F8 - 196.00 312.66 3.9e-09 45678  Z70719 Human DNA sequence fr
gb.pr:BC003190 + 190.00 337.75 1.9e-10 793  BC003190 Homo sapiens, p75NTR-
gb.pr:HSU80GC + 190.00 336.71 1.8e-10 891  M38188 Human unknown protein fr
gb.pr:AF187064 + 190.00 336.71 1.8e-10 891  AF187064 Homo sapiens p75NTR-
gb.pr:AK000959 + 187.00 328.38 5.2e-10 1329  AK000959 Homo sapiens cDNA FLJ
gb.pat:AX100231 + 187.00 327.45 5.8e-10 1364  AX100231 Sequence 32 from Pat
gb.pr:HS635G19 + 187.00 292.48 5.1e-08 69648  AL035494 Human DNA sequence
gb.pr:AF187066 + 183.50 327.00 6.2e-10 700  AF187066 Mus musculus p75NTR-
gb.ro:AF097440 + 183.50 325.23 7.7e-10 854  AF097440 Mus musculus brain exp
gb.sts:G35294 - 170.00 305.77 9.4e-09 477  G35294 human STS SHGC-37409, se
gb.ro:AF187065 + 150.50 269.43 9.9e-07 519  AF187065 Rattus norvegicus p75N
gb.sts:G72708 - 140.50 253.04 8.1e-06 421  G72708 MARC 4953-4954:991939031
gb.sts:G23964 - 103.50 186.62 0.0406 372  G23964 human STS WI-15922, sequ
gb.om:OATPICH - 93.50 139.71 16.63 9344  Z18361 O.aries trichohyalin.2
gb.ba:SCHE69 - 93.50 127.76 77.00 35924  AL079308 Streptomyces coelic
gb.pr:AB040971 + 92.50 144.57 8.92 4408  AB040971 Homo sapiens mRNA for
gb.hlg:AC068585 + 92.50 119.94 210.09 70358  AC068585 Homo sapiens clone H
gb.hlg:AF003575 + 90.00 109.80 771.39 146017  AF003575 Oryza sativa Chrom
gb.om:OCTRICHA - 88.00 133.54 36.71 9121  Z19092 O.cuniculus trichohyal
gb.ba:AF010496 - 88.00 102.92 1.9e+03 183370  AF010496 Rhodobacter capsula
gb.ro:AB041591 - 87.50 144.14 9.43 1657  AB041591 Mus musculus brain c
gb.pat:AX188346 + 87.00 145.40 8.02 1298  AX188346 Sequence 4041 from Pa
gb.pat:AF64152 + 87.00 141.54 13.16 2004  AF64152 Sequence 18 from Patent
gb.pat:AF68338 + 87.00 141.54 13.16 2004  AF68338 Sequence 18 from Patent
gb.pat:AR031270 + 87.00 141.54 13.16 2004  AR031270 Sequence 18 from pat
```

```
gb_pat:AR036392 + 87.00 141.54 13.16 2004  AR036392 Sequence 18 from I
gb_pat:AR060824 + 87.00 141.54 13.16 2004  AR060824 Sequence 18 from I
gb_pat:AR065979 + 87.00 141.54 13.16 2004  AR065979 Sequence 18 from I
gb_pat:AR089933 + 87.00 141.54 13.16 2004  AR089933 Sequence 18 from I
gb_pat:AF64169 + 87.00 139.30 17.52 2576  AF64169 Sequence 35 from Pat
seq_name: gb_pr:AF251053
seq_documentation_block:
LOCUS AF251053 744 bp mRNA PRI 15-APR-2001
DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
ACCESSION AF251053
VERSION AF251053.1 GI:13625167
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 744)
AUTHORS Mao,Y., Xie,Y., Zhou,Z., Zhao,W., Zhao,S., Wang,W., Huang,Y.,
Wang,S., Tang,R., Chen,X. and Wu,C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life
Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.
China
FEATURES
source Location/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
101..487
/codon_start=1
/product="X-linked protein"
/protein_id="AAK34943.1"
/db_xref="GI:13625168"
/translatability="MESKERALNNLIVNVOENDEKDEQVANKGEPLALPLNVS
EYCVPRGNRRFRVROPILQYRWDIMHRLGEPQARREMERIEEVRQLMKLREK
OLSHSLRAVSTDPHHDHDFCLMP"
BASE COUNT 222 a 134 c 208 g 180 t
ORIGIN
alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-31 x AF251053
Align seg 1/1 to: AF251053 from: 1 to: 744
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
101 ATGGAGTCCAAAGAACGAGCGGTAAACAATCTCATCGTGGAAATGT 150
17 LasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaLeuLysG 34
151 CAACAGGAAATATGATAAAGATGAAAGGAGCAAGTCTGCTATATAAG 200
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
201 GGGAGCCCTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCTAGA 250
51 GlyAsnArgArgArgheArgValArgGlnProLeuGlnTyrArgTr 67
251 GGAACCGTAGGCGGTTCGCGGTAGGAGCCATCTCGCATATAGATG 300
67 pasPileMetHisArgLysGluGluProGlnAlaArgMetArgGluGluA 84
301 GGACATAATCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 350
84 snMetGluArgLysGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
```

351 ATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 400
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
|||||
401 GAAAGCAGATTGAGTATAGTCTCGGGCAGTCTGAGCACTGATCCCCCTCA 450
117 shisAspHisHisAspGluPheCysLeuMetPro 128
|||||
451 CCATGACCATCAGATGAGTTTGGCTTATGCCC 484

seq_name: gb_pat:AX078272

seq_documentation_block: 898 bp DNA PAT 22-FEB-2001
LOCUS AX078272
DEFINITION Sequence 76 from Patent WO0107471.
ACCESSION AX078272
VERSION AX078272.1 GI:13157963
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 898)
AUTHORS Hillman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Asimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 76 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..898
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 5664154CB1"

BASE COUNT 250 a 186 c 251 g 211 t
ORIGIN
alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x AX078272

Align seg 1/1 to: AX078272 from: 1 to: 898

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
254 ATGGAGTCCAAAGAGGAGGAGGCTTAACAATCTCATGCTGGAATAATGT 303
17 LasnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
|||||
304 CAACCCGGAATAATGATGAAGAAGATGAAGAAGCAAGTGTCTAATAAG 353
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
354 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATACCTGTGCTGCTAGA 403
51 GlyAsnArgArgGluGluValArgGlnProIleLeuGlnTyrArgTr 67
|||||
404 GGAACCCGTAGCGGTTCCCGTTAGGAGCCCATCTCTGCATGTATAGTG 453
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
|||||
454 GGACATAATGATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 503
84 snMetGluArgGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
504 ATATGGAAGGATTGGGAGAGGAGTGAGACAGCTGTATGGAAGAGCTGAGG 553
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
|||||

554 GAAAGCAGCTTGGTATAGTCTCGGGCAGCTCAGCACTGATCCCCCTCA 603

117 shisAspHisHisAspGluPheCysLeuMetPro 128
|||||
604 CCATGACCATCAGATGAGTTTGGCTTATGCCC 637

seq_name: gb_pr:HSV870H8

seq_documentation_block:

LOCUS HSV870H8 31321 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid:V870H8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.

ACCESSION 270233

VERSION 270233.1 GI:1235542

KEYWORDS X.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 31321)

AUTHORS Whitelev,M.

TITLE Direct Submission

JOURNAL Submitted (19-MAR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT
IMPORTANT: This sequence is not the entire insert of clone V870H8.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.

The true left end of clone V870H8 is at 1 in this sequence. The
true left end of clone V693A8 is at 31221.

V870H8 is from the human chromosome X-specific cosmid library.

FEATURES

source

1..31321

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="X"

/clone="CHC-870H8"

/clone_lib="SCCV"

475..701

/note="L1 element fragment"

802..988

/note="L1 element fragment"

2533..4138

/note="match: multiple ESTs"

4189..4246

/note="29 copies of 2 mer 91 % conserved"

7496..7874

/note="L1 element fragment"

8022..8108

/note="L1 element fragment"

8302..8493

/note="L1 element fragment"

8585..8671

/note="L1 element fragment"

8769..9010

/note="L1 element fragment"

9259..9384

/note="L1 element fragment"

9624..9731

/note="MLTIC element fragment"

9781..9948

/note="MLTIC element fragment"

9890..9948

/note="MLTIB element fragment"

10418..10592

/partial

/note="Alu repeat: matches 298..108 of consensus"

10593..10688

/partial

/note="Alu repeat: matches 96..1 of consensus"

10698..10781

/note="L1 element fragment"

```

repeat_region      11112..11403
repeat_region      11922..12004
repeat_region      12062..12308
repeat_region      12423..12707
repeat_region      12710..12916
repeat_region      12979..13081
repeat_region      13473..13532
repeat_region      15871..15928
repeat_region      18180..18417
repeat_region      18778..19056
repeat_region      22279..22350
repeat_region      22449..22995
repeat_region      23239..23374
repeat_region      23371..23606
repeat_region      23395..23541
repeat_region      23705..23778
repeat_region      23805..23960
repeat_region      24011..24541
repeat_region      24638..24770
repeat_region      24829..30975
BASE COUNT      9462 a 6147 c 6129 g 9593 t
ORIGIN

alignment_scores:
    Quality: 679.00      Length: 128
    Ratio: 5.305         Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x HSV870H8

Align seg 1/1 to: HSV870H8 from: 1 to: 31321

1 MetGluSerLySgLUgUArgAlaLeuAAsnAsnLeuIlleValGluAsnVa 17
3508 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCGTGGAAAATGT 3557
17 lAsnGInGluAsnAspGluLySgLUgUArgAlaLeuValAlaAsnLySg 34
3558 CAACCGAGGAATAATGATGAAGAAGATGAAAGGACGAAGTTGGTAATAAG 3607
34 lYGIuPrOLeuAlaLeuPrOLeuAsnValSerGIuTyfCysValPrOArg 50
3608 GGGAGCCCTGGCCCTACCTTTCGAATGTTAGTGAATCTGTGTGCTTAGA 3657
51 GlyAsnArgArgArgPheArgValArgInPrOIlleLeuGInTyfArgTr 67
3658 GGAACCGTATAGCGGTTCCGCGTTAGGCACGCCCATCTCCGACGTATAGATG 3707
67 pAspIleMetHisArgLeuGlyGIuPrOGInAlaArgMetArgGIuGluA 84

```



```
alignment_scores:
  Quality: 679.00      Length: 128
  Ratio: 5.305         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x AL133348/rev ..

Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
31132 ATGGAGTCCAAAGAGGACGCGTTAAACAATCTCATCGTGGAAATGT 31083
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
31082 CAACCCAGGAAATGATGAAAGAGATGAAAGAGCAAGTTCCTAATAAG 31033
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
31032 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGCTCTAGA 30983
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTr 67
30982 GGAACCGTAGCGGTTCCGGTTAGGAGCCCATCTCGCAGTATAGATG 30933
67 pAspTleMetHisArgLeuGlyGluProGlnAlaAlaArgMetArgGluA 84
30932 GGACATAATGATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 30883
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
30882 ATATGGAAGGATTTGGGAGGAGTGAGACAGCTGATGGAAGCTGAGG 30833
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
30832 GAAAAGCAGTTGAGTCACTAGTTTGGCGGAGTGCAGTGCATCCCTCA 30783
117 shIsAspHisHisAspGluPheCysLeuMetPro 128
30782 CCATGACCATCAGATGAGTTTGCCTTATGCC 30749

seq_name: gb_htg:HSU80B1

seq_documentation_block:
LOCUS      HSU80B1 41029 bp  DNA  HTG  10-JUL-2001
DEFINITION Homo sapiens chromosome X clone LLOXNC01-80B1 map q22.1-22.3, ***
ACCESSION AL022169
VERSION    AL022169.3 GI:13276704
KEYWORDS   HTG; HTGS_PHASE2; HTGS_CANCELLED.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 41029)
AUTHORS   Bird,C.
TITLE     Direct Submission
JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            requests: clonerequest@sanger.ac.uk
            On Mar 12, 2001 this sequence version replaced gi:12750927.
COMMENT   ----- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ----- Project Information
            Center project name: cu80B1
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; 108752; 100% of reads
            Chemistry: Dye-terminator ABI; 19% of reads

Chemistry: Dye-terminator Big Dye; 80% of reads
Consensus quality: 41009 bases at least Q40
Consensus quality: 41023 bases at least Q30
Consensus quality: 41027 bases at least Q20
Insert size: 41029; sum-of-contigs
Insert size: 51751; 0.6% error; agarose-fp
Quality coverage: 9.67x in Q20 bases; sum-of-contigs Quality
coverage: 7.67x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES             Location/Qualifiers
     source            1..41029
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="q22.1-22.3"
                     /clone="LLOXNC01-80B1"
                     /clone_lib="LLOXNC01"
     misc_feature      1..41029
                     /note="assembly_fragment:00890"
BASE COUNT  13249 a 7796 c 7781 g 12203 t
ORIGIN

alignment_scores:
  Quality: 679.00      Length: 128
  Ratio: 5.305         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-327-750D-31 x HSU80B1 ..

Align seg 1/1 to: HSU80B1 from: 1 to: 41029

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
18412 ATGGAGTCCAAAGAGGACGCGTTAAACAATCTCATCGTGGAAATGT 18461
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
18462 CAACCCAGGAAATGATGAAAGAGATGAAAGAGCAAGTTCCTAATAAG 18511
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
18512 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGCTCTAGA 18561
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTr 67
18562 GGAACCGTAGCGGTTCCGGTTAGGAGCCCATCTCGCAGTATAGATG 18611
67 pAspTleMetHisArgLeuGlyGluProGlnAlaAlaArgMetArgGluA 84
18612 GGACATAATCCTAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 18661
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
18662 ATATGGAAGGATTTGGGAGGAGTGAGACAGCTGATGGAAGCTGAGG 18711
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
18712 GAAAAGCAGTTGAGTCACTAGTTTGCCTTATGCC 18761
117 shIsAspHisHisAspGluPheCysLeuMetPro 128
18762 CCATGACCATCAGATGAGTTTGCCTTATGCC 18795

seq_name: gb_pr:AF183416
seq_documentation_block:
LOCUS      AF183416 642 bp  mRNA  PRI  02-SEP-2000
```

```

DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
mRNA, complete cds.
ACCESSION AF183416
VERSION AF183416.1 GI:9963770
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Peng, Y., Qian, B., Tu, Y., Xu, S., Han, Z., Fu, G. and Chen, Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE
AUTHORS Peng, Y., GU, W., Huang, C., Xu, S., Han, Z., Fu, G. and Chen, Z.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source
1..642
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
3..380
/notes="HGR74-h protein"
/codon_start=1
/evidence="not_experimental"
/product="ovarian granulosa cell 13.0 kDa protein hGR74
homolog"
1..642
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
3..380
/notes="HGR74-h protein"
/codon_start=1
/evidence="not_experimental"
/product="ovarian granulosa cell 13.0 kDa protein hGR74
homolog"
BASE COUNT 204 a 118 c 157 g 163 t
ORIGIN
alignment_scores:
Quality: 606.50 Length: 128
Ratio: 4.891 Gaps: 1
Percent Similarity: 96.875 Percent Identity: 89.062
alignment_block:
US-09-327-750D-31 x AF183416
Align seg 1/1 to: AF183416 from: 1 to: 642
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuValGluAsnVa 17
|||||
3 ATGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 52
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysg 34
|||||
53 CAACCAAGAGAAAT.....GAAGAAAGAGAGAGAGAGAGAGAG 93
|||||
34 lYGlupProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
94 GGAGAGCCCTTGCCCTCCCTTTGGATGCTGGTGAATCTGTGTCCTAGA 143
|||||
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTtr 67
|||||
144 GGAATCGTAGGCGGTTCGCGCTTAGCGAGCCCATCTCGAGTATAGATG 193
|||||
67 pAspIleMetHisArgLeuGluProGlnAlaArgMetArgGluAla 84
|||||
194 GGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 243
|||||
84 snMetGluArgTleGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
244 ATATGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGAAAGAGCTGAGG 293
|||||

```

```

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
294 GAAAGACAGTTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCA 343
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
344 CCATGACCATCATGATGAGTTTTCCTTATGCC 377
|||||
seq_name: gb_pr:AF220189
seq_documentation_block: 828 bp mRNA PRI 04-MAY-2000
LOCUS AF220189
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189
VERSION AF220189.1 GI:7689028
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Xiao, H., Song, H., Gao, G., Ren, S., Chen, Z. and Han, Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
REFERENCE
AUTHORS Xiao, H., Song, H., Gao, G., Ren, S., Chen, Z. and Han, Z.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67654.1"
/db_xref="GI:7689028"
/translation="MESKEKRAVNSLSMENAOEKEQVANKGEPLALPLDAGEYC
VPGNRFRFRVROPILQYRWDMHRLGEPQARMREENMERIGEVRQLMEKLRKQLS
HSLRAVSTDPHHDHDECLMP"
BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN
alignment_scores:
Quality: 606.50 Length: 128
Ratio: 4.891 Gaps: 1
Percent Similarity: 96.875 Percent Identity: 89.062
alignment_block:
US-09-327-750D-31 x AF220189
Align seg 1/1 to: AF220189 from: 1 to: 828
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuValGluAsnVa 17
|||||
167 ATGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysg 34
|||||
217 CAACCAAGAGAAAT.....GAAGAAAGAGAGAGAGAGAGAGAG 257
|||||
34 lYGlupProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
258 GGAGAGCCCTTGCCCTCCCTTTGGATGCTGGTGAATCTGTGTCCTAGA 307
|||||
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTtr 67
|||||
308 GAAATCGTAGGCGGTTCGCGCTTAGCGAGCCCATCTCGAGTATAGATG 357
|||||

```

```

67 pAspilewHisArgLeuGlyProGlnAlaArgMetArgGluGlnA 84
||||:|||||
358 GGATATGATCATAGCTGGAGACCAAGGAGGATGAGAGAGAGA 407
|||||
84 snMetGluArgIleGlyGluGlnValArgGlnLeuMetGluLysLeuArg 100
|||||
408 ATATGGAAGAGATTGGGAGGAGGCTGAGACAGCTGATGAAAGCTGAGG 457
|||||
101 GluLysGlnLeuSerHisLeuArgAlaValSerThrAspProH1 117
|||||
458 GAAAGCAGTGTAGTCTGCGGGCAGTCAGCACTGACCCCTCA 507
|||||
117 sHisAspHisAspGluPheCysLeuMetPro 128
|||||
508 CCATGACCATCATGATGAGTTTGGCTTATGCC 541
|||||

```

seq_name: gb_pr:HS198P4

seq_documentation_block: 16-DEC-2000
LOCUS HS198P4 35714 bp DNA PRI
DEFINITION Human DNA sequence from clone RP1-198P4 on chromosome Xq22 Contains
a CpG Island, complete sequence.

ACCESSION AL008708
VERSION AL008708.4 GI:4826445
KEYWORDS HTG; CpG Island.
SOURCE human.

ORGANISM

Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Corby, N.

REFERENCE

AUTHORS
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT

requests: clonerequest@sanger.ac.uk
On May 13, 1999 this sequence version replaced gi:4582117.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep
from the library RPI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone
RPI-198P4. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

The true left end of clone LLOXNC01-221F2 is at 35615 in this
sequence. The true right end of clone LLOXNC01-73E8 is at 100 in
this sequence.

FEATURES

Location/Qualifiers

1..35714
/organism="Homo sapiens"
/db_xref="taxon:9606"

/chromosome="X"
/map="q22"
/clone="RP1-198P4"
/clone_lib="RPC1-1"
1..258
/note="L1PA3 repeat: matches 5666..5925 of consensus"
256..537
/note="L1PA3 repeat: matches 5391..5672 of consensus"
1491..1749
/note="AluSc repeat: matches 39..297 of consensus"
2182..3810
/note="L1PB3 repeat: matches 4533..6150 of consensus"
5169..6687
/note="L1MA2 repeat: matches 4798..6308 of consensus"
6976..7665
/note="L1MA2 repeat: matches 4113..4798 of consensus"
8039..9213
/note="L1MA2 repeat: matches 2945..4113 of consensus"
9523..10437
/note="L1MA2 repeat: matches 2016..2945 of consensus"
10736..11494
/note="L1MA2 repeat: matches 1262..2016 of consensus"
11807..12251
/note="L1MCA repeat: matches 310..755 of consensus"
12259..12331
/note="HERVH21 repeat: matches 5..77 of consensus"
12332..12796
/note="L1R21A repeat: matches 1..490 of consensus"
16162..16371
/note="105 copies 2 mer tt 55% conserved"
17121..17424
/note="MER33 repeat: matches 1..322 of consensus"
18259..18290
/note="16 copies 2 mer ca 87% conserved"
18505..18672
/note="MER63A repeat: matches 1..181 of consensus"
19478..19814
/note="MER2 repeat: matches 1..345 of consensus"
21061..21101
/note="HERVH21 repeat: matches 42..85 of consensus"
21247..21781
/note="MLT1E repeat: matches 44..568 of consensus"
24465..24492
/note="Single clone region"
Single clone region
24531..24562
/note="Weak data"
Weak data
24563
/note="Single clone region"
Single clone region
24711..24865
/note="Region covered by subclones from a PAC DNA PCR
product only at 12x coverage."
Region covered by subclones from a PAC DNA PCR product only
at 12x coverage.
24758..25460
/evidence="not_experimental"
24875..24940
/note="33 copies 2 mer cc 66% conserved"
26795..26926
/note="66 copies 2 mer ct 63% conserved"
27088..27133
/note="23 copies 2 mer tg 100% conserved"
27461..27777
/note="L1MB8 repeat: matches 5851..6171 of consensus"
27802..28232
/note="L1MCD repeat: matches 5192..5403 of consensus"
28536..28930
/note="L1MCD repeat: matches 4802..5192 of consensus"
29240..29341
/note="TIGER1 repeat: matches 484..587 of consensus"


```
repeat_region 29603..30740
/Note="RIGER1 repeat: matches 587..1757 of consensus"
repeat_region 31027..31365
/Note="RIGER1 repeat: matches 1757..2105 of consensus"
repeat_region 31664..31691
/Note="RIGER1 repeat: matches 2105..2132 of consensus"
repeat_region 31809..32015
/Note="RIGER1 repeat: matches 2132..2246 of consensus"
repeat_region 32323..32389
/Note="RIGER1 repeat: matches 2246..2313 of consensus"
repeat_region 34353..34636
/Note="MLT1-INTERNAL repeat: matches 607..889 of consensus"
repeat_region 35149..35376
/Note="MLT1-INTERNAL repeat: matches 1001..1233 of consensus"
repeat_region 35523..35670
/Note="MLT1C repeat: matches 8..159 of consensus"
BASE COUNT 9708 a 7094 c 6805 g 12107 t
ORIGIN
alignment_scores:
  Quality: 606.50 Length: 128
  Ratio: 4.891 Gaps: 1
  Percent Similarity: 96.875 Percent Identity: 89.062
alignment_block:
US-09-327-750D-31 x HS198P4/rev
Align seg 1/1 to reverse of: HS198P4 from: 1 to: 35714
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
24358 ATGGAGTCCAAAGAGAACGAGCAGTAACAGCTCTCAGCATGGAATGC 24309
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
24308 CAACCAAGAAAT.....GAAGAAAGAGGAGCAAGTTGCTAATAAG 24268
|||||
34 lYgluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
24267 GGGAGCCCTTGGCCCTCCCTTGGATGCTGGTGAATAGTGTGCTAGA 24218
|||||
51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
24217 GGAATTCGTAGCGGTTCCGGTTAGCAGCCATCTCGCAGTATAGATG 24168
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
24167 GGATATGATGATAGCTGGAGGAGGTGAGCAGCTGATGGAAGAGAGA 24118
|||||
84 snMetGluArgIleGlyGluGlnValArgGlnLeuMetGluLysLeuArg 100
|||||
24117 ATATGGAAGGATTGGGAGGAGGTGAGCAGCTGATGGAAGAGCTGAGG 24068
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
24067 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCAGTACGCCCTCA 24018
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
24017 CCATGACCATCATGATGAGTTTTCCTTATGCC 23984
|||||
seq_name: gb_pr:AF237783
seq_documentation_block:
LOCUS AF237783 791 bp mRNA PRI 02-SEP-2000
DEFINITION Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete cds.
ACCESSION AF237783
VERSION AF237783.1 GI:9963898
KEYWORDS
```

```
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 791)
AUTHORS Yang,Q.S., Ying,K., Xie,Y. and Mao,Y.M.
TITLE A Novel Human X-linked Brain Expressed Protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 791)
AUTHORS Mao,Y.M., Xie,Y., Yang,Q.S., Wu,H., Lin,S. and Ying,K.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2000) Genetic Research Institute, Fudan
University, 220 Handan Rd., Shanghai 200433, P.R.China
FEATURES
source 1..791
location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="X"
  /map="between DXS990 and DX1059 by RH (Stanford G3)"
  1..791
  /gene="BEX1"
  /CDS
    172..549
    /codon_start=1
    /product="brain-expressed protein BEX1"
    /protein_id="AAG09752.1"
    /db_xref="GI:9963898"
    /translation="MESKEKLAVNSLSMENANOEKEQVANKGEPLALPLDAGEVC
    VPGNRFRFRVQPIQYRWDMHRLGEPQARMREENMERIGESVRLMEKLEKQLS
    HSLRAVSTDPHHDHDEFCLMP"
BASE COUNT 214 a 172 c 218 g 187 t
ORIGIN
alignment_scores:
  Quality: 592.50 Length: 128
  Ratio: 4.857 Gaps: 1
  Percent Similarity: 95.312 Percent Identity: 87.500
alignment_block:
US-09-327-750D-31 x AF237783
Align seg 1/1 to: AF237783 from: 1 to: 791
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
172 ATGGAGTCCAAAGAGAACTAGCAGTAACAGCTCTCAGCATGGAATGC 221
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
222 CAACCAAGAAAT.....GAAGAAAGAGGAGCAAGTTGCTAATAAG 262
|||||
34 lYgluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
263 GGGAGCCCTTGGCCCTCCCTTGGATGCTGGTGAATAGTGTGCTAGA 312
|||||
51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
313 GGAATTCGTAGCGGTTCCGGTTAGCAGCCATCTCGCAGTATAGATG 362
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
363 GGATATGATGATAGCTGGAGGAGGTGAGCAGCTGATGGAAGAGAGA 412
|||||
84 snMetGluArgIleGlyGluGlnValArgGlnLeuMetGluLysLeuArg 100
|||||
413 ATATGGAAGGATTGGGAGGAGGTGAGCAGCTGATGGAAGAGCTGAGG 462
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
463 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCAGTACGCCCTCA 512
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
```

|||||
513 CCATGACCATCATGATGAGTTTGCCTTATGCC 546

seq_name: gb_ro:AF097439

seq_documentation_block: 785 bp mRNA 13-APR-1999
LOCUS AF097439 Mus musculus brain expressed X-linked protein 2 (Bex2) mRNA,
DEFINITION complete cds.
ACCESSION AF097439
VERSION AF097439.1 GI:4580591
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS 1 (bases 1 to 785)
TITLE Brown, A.L. and Kay, G.F.
JOURNAL Bex1, a gene with increased expression in parthenogenetic embryos,
MEDLINE Hum. Mol. Genet. 8 (4), 611-619 (1999)
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May; 8(5):943]]
REFERENCE 2 (bases 1 to 785)
AUTHORS Brown, A.L. and Kay, G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
source
1. 785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:W48832"
/chromosome="X"
/map="near Pip"
/dev_stage="embryo; 15.5 dpc"
gene 1. 785
/gene="Bex2"
CDS 139..528
/gene="Bex2"
/codon_start=1
/product="brain expressed X-linked protein 2"
/protein_id="A024430.1"
/db_xref="GI:4580592"
/translation="MESKVEQGVKNLMNENDHOKKEEKPDASKRDPTVALPFEA
GDYVPRGRRFRVROPVIVHVRWDLMDLRVGEPOGRMRREENVQFGDDVRLMEKLRER
ROLSHSLRAVSTDPHHDHDFCLMP"

BASE COUNT 212 a 163 c 235 g 175 t

ORIGIN

alignment_scores: Quality: 502.50 Length: 129

Ratio: 4.153 Gaps: 1

Percent Similarity: 93.798 Percent Identity: 71.318

alignment_block:

US-09-327-750d-31 x AF097439

Align seg 1/1 to: AF097439 from: 1 to: 785

1 MetGlusLysGluGluArgAlaLeuAsnLeuIleValGluAsnVa 17

|||||
139 ATGAGTCCTCAAGTGAACACAGCGTGAATAATCTCAACATGGAGAATGA 188

17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

|||||
189 CCATCAGGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 238

34 IyGluProLeu...AlaLeuProLeuAsnValSerGluTyrCysValPro 49

|||||
:|||||: |||||||:|||||:|||||:|||||

239 GGGATCCGATTGTGGCCCTCCCTTTTCGAAGCTGGAGACTACTACGTGCCT 288
50 ArgGlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66
|||||
289 AGAGGAGTGCAGCGGTTCCGGGTTCCGCAGCCCATCGTGCACACTACAG 338
66 gTTPAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluG 83
|||||
339 ATGGACCTGATGCTAGGTTGGGAGCCCGCCAGGAGGAGGAGGAGGAGG 388
83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
|||||
389 AGAACGTACAGAGGTTGGGATGATGTGAGACAGCTCATGAGAAGCTG 438
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||
439 AGGAAGGAGCTGAGCCACAGCTGCGGGGTTAGCACTGACGCCGCC 488
116 OHSHISAPHSIHISASpGluPheCysLeuMetPro 128
|||||
489 TCATCATGACCATCATGATGAGTTTGCCTTATGCC 525

seq_name: gb_ro:AF097437

seq_documentation_block: 2269 bp DNA 05-AUG-1999
LOCUS AF097437 Mus musculus Bex1 protein (Bex1) gene, complete cds.
DEFINITION
ACCESSION AF097437
VERSION AF097437.1 GI:5702155
KEYWORDS house mouse.
SOURCE
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS Brown, A.L. and Kay, G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
PUBMED 10072429
REFERENCE 2 (bases 1 to 2269)
AUTHORS Brown, A.L. and Kay, G.F.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4030, Australia
FEATURES
Location/Qualifiers
1. 2269
/organism="Mus musculus"
/strain="129SVJ"
/db_xref="taxon:10090"
/chromosome="X"
/map="near Pip"
/tissue_type="liver"
join(409..509,866..942,1263..1873)
/gene="Bex1"
/product="Bex1 protein"
409..1873
/gene="Bex1"
/note="expressed in brain; X-linked"
1268..1654
/gene="Bex1"
/codon_start=1
/product="Bex1 protein"
/protein_id="A047168.1"
/db_xref="GI:5702156"
/translation="MESKDGKVNLMNENDHOKKEEKPDQIRREPAVALTSEAG
KNCAPGRRFRVROPVIAHVRWDLMDLRVGEPOGRMRREENVQFGDDVRLMEKLRER
QLSHSLRAVSTDPHHDHDFCLMP"

BASE COUNT 563 a 554 c 705 g 447 t

ORIGIN

alignment_scores:

Quality: 452.00 Length: 129
Ratio: 3.897 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:

US-09-327-750D-31 x AF097437

Align seg 1/1 to: AF097437 from: 1 to: 2269

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
1268 ATGGAGTCCAAAGATCAA...GGCGTGAATAATCTCAACATGGAGATGA 1314
17 LasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
1315 CCATCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1364
34 lyGluPro...LeuAlaLeuProLeuAsnValSerGluTyrCysValPro 49
1365 GGGAGCCAGCTGTGCGCTGACCTCGAGGCTGGCAAACTGTGCACCT 1414
50 ArgGlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66
1415 AGAGGAGGTCGAGCGGCTTCCGGGTTCGCGAGCCATCGCTCCTATAG 1464
66 gTrpAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluG 83
1465 ATGGAGCTGTATGAGAGGTTGGGAGCCCGGAGGAGGAGGAGGAGGAGG 1514
83 LuAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
1515 AGAAGCTACAGAGGTTTGGGGGTGTGTGAGCAGCTCATGGAGAGCTG 1564
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
1565 AGGGAAGGAGCTGTGAGCAGCTCGGGGGGTGTGAGCAGCTCATGCC 1614
116 ohHisAspHisHisAspGluPheCysLeuMetPro 128
1615 TCATCATGACACCATGATGAGTTTGCCTCATGCC 1651

seq_name: gb_ro:AF097438

seq_documentation_block: 835 bp mRNA ROD 11-APR-1999
LOCUS AF097438 Mus musculus brain expressed x-linked protein 1 (Bex1) mRNA,
DEFINITION complete cds.
ACCESSION AF097438
VERSION AF097438.1 GI:4580589
KEYWORDS house mouse,
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Brown,A.L. and Kay,G.F.
AUTHORS Bex1, a gene with increased expression in parthenogenetic embryos,
TITLE is a member of a novel gene family on the mouse x chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE 2 (bases 1 to 835)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
source 1..835
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"

/map="near Plp"
/tissue_type="blastocyst"
1..835
/gene="Bex1"
205..591
/gene="Bex1"
/codon_start=1
/product="brain expressed x-linked protein 1"
/protein_id="AAD24429.1"
/db_xref="GI:4580590"

gene

CDS

BASE COUNT 238 a 184 c 246 g 167 t
ORIGIN
alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667
alignment_block:
US-09-327-750D-31 x AF097438
Align seg 1/1 to: AF097438 from: 1 to: 835

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
205 ATGGAGTCCAAAGATCAA...GGCGTGAATAATCTCAACATGGAGATGA 251
17 LasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
252 CCATCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301
34 lyGluPro...LeuAlaLeuProLeuAsnValSerGluTyrCysValPro 49
302 GGGAGCCAGCTGTGCGCTGATCTCCGAGGCTGGCAAACTGTGCGCT 351
50 ArgGlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66
352 AGAGGAGGTCGAGCGGTTCCGGGTTCGCGAGCCATCGCTCCTATAG 401
66 gTrpAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluG 83
402 ATGGAGCTGTATGAGAGGTTGGGAGCCCGGAGGAGGAGGAGGAGGAGG 451
83 LuAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
452 AGAAGCTACAGAGGTTTGGGGGTGTGTGAGCAGCTCATGGAGAGCTG 501
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
502 AGGGAAGGAGCTGTGAGCAGCTCGGGGGGTGTGAGCAGCTCATGCC 551
116 ohHisAspHisHisAspGluPheCysLeuMetPro 128
552 TCATCATGACACCATGATGAGTTTGCCTCATGCC 588

seq_name: gb_ro:AF051347

seq_documentation_block: 838 bp mRNA ROD 01-OCT-1998
LOCUS AF051347
DEFINITION Mus musculus REX-3 mRNA, complete cds.
ACCESSION AF051347
VERSION AF051347.1 GI:3510642
KEYWORDS house mouse,
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 838)
AUTHORS Faria,T.N., Larosa,G.J., Wilen,E., Liao,J. and Gudas,L.J.

FEATURES

	BASE COUNT
	ORIGIN

alignment_block:

Align seg 1/1 -to: AF051347

```

1..2901
location/Qualifiers
organism="Mus musculus"
gb_xref accession="10090"
clone="MGC:358 IMAGE:3500522"
title="C3(1) Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma, 3 month old virgin mouse."
clone_id="MGC_C3(1)_GAP_Mam0"
lab_host="BALB/c"

```

CDS

alignment_scores:

seq_name: gb_ro:BC003254

seq_documentation_block:

Align seg 1/1 to: BC003254 from: 1 to: 2901

```

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 ATGGAGTCCAAAGATCAA...GGCGTGAATAATCTCAACATGGAGAATCA 269
17 laenGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 CCATCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 319
34 lyGluPro...LeuAlaLeuProLeuAsnValSerGluTyrCysValPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 GGGAGCCACGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACGTGCACT 369
50 ArgGlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AGAGGAGGTCTCGAGCGGTTCCGGGTTCCGCGAGCCCATCGCTACTATAG 419
66 gTrpAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 ATGGACCTGTATGTCAGAGGGTTGGGAGCCGCCAGGAGGAGGAGGAGG 469
83 luAsnMetGluArgIleGluGluValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 AGAACGTACAGAGGTTGGGGTGTATGTGAGACAGCTCATGGAGAGCTG 519
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 AGGGAAGGAGCTGAGCCAGCTCGGGCGGTAGCACTAGCCCGCC 569
116 oHisHis 118
|||||
570 TCATCAT 576

```

seq_name: gb_sts:G24641

seq_documentation_block: 504 bp DNA STS 31-MAY-1996
 LOCUS G24641
 DEFINITION human STS WI-11354, sequence tagged site.
 ACCESSION G24641
 VERSION G24641.1 GI:1344967

KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE human STSs derived from sequences in dbEST and the Unigene collection.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 504)
 AUTHORS Hudson,T.
 TITLE Whitehead Institute/MIT Center for Genome Research; Physically
 Mapped STSS
 JOURNAL Unpublished (1995)

COMMENT

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: TTTTTCGCTTTTAAAGGTGC
 Primer B: TTACGTGGGTCTCCTATTACCA
 STS size: 150
 PCR Profile:

Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng

Primer: each 5 pM
 dNTPs: each 4 nM
 Taq Polymerase: 0.025 units/ul
 Total Vol: 20 ul

Buffer:
 MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCL: 10 mM
 pH: 9.3

Derived from dbEST (genbank accession T86927).

FEATURES
 source 1.504
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 STS
 primer_bind 1..150
 primer_bind 1..22
 primer_bind complement(129..150)
 BASE COUNT 134 a 118 c 101 g 144 t 7 others
 ORIGIN

alignment_scores:
 Quality: 339.00 Length: 85
 Ratio: 4.644 Gaps: 3
 Percent Similarity: 85.882 Percent Identity: 83.529

alignment_block:
 US-09-327-750D-31 x G24641/rev ..

Align seg 1/1 to reverse of: G24641 from: 1 to: 504

```

47 CysValProArgGly.AsnArgArgArgPhe.ArgValArgGlnProIle 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
503 TGTCCCCCTANAGGAANCCGTAGCGGTTCCCCCGTTAGGCANCCCCAC 454
63 LeuGlnTyrArgTrpAspIleMetHisArgLeuGlyGlu.ProGlnAlaA 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
453 CCGCAGTANAGATGGGATATGATGCATAGGCTTGGAGAACCACAGCAA 404
79 rgMetArgGluGluAsnMetGluArgIleGlyGluGluValArgGlnLeu 95
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
403 GGATGAGAGAAAGAGATATGGAAGGATTGGGAGGAGGTGAGACACCTG 354
96 MetGluLysLeuArgGluLysGlnLeuSerHisSerLeuArgAlaValSe 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 ATGGAAGAGCTGAGGGAAGACAGTTGAGTCANAGTCTGGCGGCGAGTCAG 304
112 rThrAspProHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 CACTGACCCCCCTCACCATGACCATCATGATGAGTTTGTGTTTATGCC 255

```

OM of: US-09-327-750d-32 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framed-p2n.model -DEV-xlh
-Q=cn22_1/USPFO_spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODEL=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750 -CGN1_1.5654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-32

Query length: 125

Database: EST*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_est1:BG655001	+	653.00	1221.08	547	1 BG655001 ib44f06.y1 HR85 islet
gb_est2:BG6472167	+	653.00	1219.74	8.9e-59	634 1 BG472167 602513894F1 NIH_MGC_16
gb_est3:BF971303	+	653.00	1219.40	9.3e-59	658 1 BF971303 602273150F1 NIH_MGC_84
gb_est4:BG714823	+	653.00	1219.18	9.6e-59	674 1 BG714823 602677146F1 NIH_MGC_98
gb_est1:AV702285	+	653.00	1218.97	9.8e-59	690 1 AV702285 AV702285 ADB Homo sapi
gb_est1:AV705808	+	653.00	1218.96	9.9e-59	691 1 AV705808 AV705808 ADB Homo sapi
gb_est1:AL526209	+	653.00	1218.91	9.9e-59	695 1 AL526209 AL526209 LTI_NFL003.NE
gb_est2:BG6820179	+	653.00	1218.85	1.0e-58	699 1 BG6820179 602782222F1 NCI_CGAP_E
gb_est1:BG790774	+	653.00	1217.82	1.1e-58	783 1 BG790774 601581956F1 NIH_MGC_7
gb_est1:AL526247	+	653.00	1217.70	1.2e-58	794 1 AL526247 AV702643 ADB Homo sapi
gb_est1:AV702643	+	653.00	1217.43	2.0e-58	683 1 AV702643 AV702643 ADB Homo sapi
gb_est1:AI929703	+	646.00	1207.46	4.3e-58	576 1 AI929703 au63f04.y1 Schneider
gb_est2:BG715659	+	646.00	1205.26	5.7e-58	734 1 BG715659 602676924F1 NIH_MGC_98
gb_est2:BG709325	+	637.00	1187.25	5.8e-57	828 1 BG709325 602673345F1 NIH_MGC_98
gb_est1:AW160832	+	631.00	1180.01	1.5e-56	531 1 AW160832 au76b12.y1 Schneider
gb_est2:BG346478	+	630.00	1176.22	2.4e-56	656 1 BG346478 60202342F1 NCI_CGAP_E
gb_est1:AV727777	+	627.00	1171.87	4.1e-56	569 1 AV727777 AV727777 HTC Homo sapi
gb_est2:AW953922	+	627.00	1171.28	4.5e-56	607 1 AW953922 EST365887 MAGE resequ
gb_est2:BG704861	+	627.00	1168.97	6.0e-56	783 1 BG704861 602883214F1 NIH_MGC_95
gb_est1:AV728527	+	626.00	1164.93	1.0e-55	993 1 AV728527 AV728527 HTC Homo sapi
gb_est2:BF966650	+	625.00	1165.09	9.9e-56	794 1 BF966650 602286982F1 NIH_MGC_95
gb_est2:BG714687	+	607.00	1132.41	6.5e-54	700 1 BG714687 602677046F1 NIH_MGC_98
gb_est1:BG314909	+	602.50	1125.45	1.6e-53	594 1 BG314909 601140074F1 NIH_MGC_9
gb_est1:AL520931	+	602.50	1123.74	2.0e-53	717 1 AL520931 AL520931 LTI_NFL004.NE
gb_est2:BE889313	+	602.50	1123.07	2.2e-53	772 1 BE889313 601513201F1 NIH_MGC_71
gb_est2:BG709503	+	602.50	1122.63	2.3e-53	810 1 BG709503 602674781F1 NIH_MGC_98
gb_est2:BG703798	+	602.50	1122.35	2.4e-53	835 1 BG703798 602672810F1 NIH_MGC_98
gb_est2:BF038787	+	602.50	1120.34	3.1e-53	1042 1 BF038787 601462119F1 NIH_MGC_98
gb_est2:BG285825	+	600.50	1117.46	4.4e-53	946 1 BG285825 602380917F1 NIH_MGC_93
gb_est2:BG036775	+	600.00	1118.04	4.1e-53	801 1 BG036775 602286982F1 NIH_MGC_95
gb_est2:CI15948	+	598.00	1118.78	3.7e-53	488 1 CI15948 CI15948 Clontech human
gb_est2:BF348593	+	598.00	1113.14	7.7e-53	908 1 BF348593 602019889F1 NCI_CGAP_E
gb_est1:AW024587	+	597.00	1115.16	6.0e-53	591 1 AW024587 wu77g08.x1 NCI_CGAP_K4
gb_est2:BG714974	+	595.50	1110.47	1.1e-52	726 1 BG714974 602677263F1 NIH_MGC_98
gb_est2:BF967675	+	595.50	1109.67	1.2e-52	793 1 BF967675 602287736F1 NIH_MGC_98
gb_est2:BG707734	+	595.50	1109.40	1.2e-52	817 1 BG707734 602671255F1 NIH_MGC_98
gb_est1:AA160657	+	594.00	1108.81	1.3e-52	639 1 AA160657 zq49f12.r1 Stratagene
gb_est2:W48778	+	592.00	1107.91	1.5e-52	467 1 W48778 zc44c03.r1 Soares_senes
gb_est2:BG705843	+	591.50	1101.93	3.3e-52	813 1 BG705843 602669329F1 NIH_MGC_98
gb_est1:AI291270	+	589.50	1102.39	3.1e-52	511 1 AI291270 qml6e06.x1 NCI_CGAP_LU

gb_est1:AI291126 + 587.50 1096.19 6.8e-52 669 1 AI291126 qml5f02.x1 NCI_CGAP
gb_est2:BF237433 + 587.50 1095.63 7.3e-52 711 1 BF237433 601842108F1 NIH_MGC
gb_est1:BE266012 + 586.50 1095.14 7.8e-52 610 1 BE266012 601194731F1 NIH_MGC
gb_est1:BE733217 + 584.50 1087.88 2.0e-51 897 1 BE733217 601568432F1 NIH_MGC
gb_est2:BF967457 + 581.50 1083.65 3.4e-51 769 1 BF967457 602287383F1 NIH_MGC

seq_name: gb_est2:BG655001

seq_documentation_block: 547 bp mRNA EST 05-JUL-2001
LOCUS BG655001
DEFINITION 1844f06.y1 HR85 islet Homo sapiens cDNA 5' similar to FR:QWRT8
QWRT8 BRAIN EXPRESSED X-LINKED PROTEIN 2.; mRNA sequence.

ACCESSION BG655001

VERSION BG655001.1 GI:13792410

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 547)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas

, M., Gibbons, M., McGann, R., Cole, R., Tsagarishvili, R., Williams, T.

, Jackson, J., and Bowers, J.

Endocrine Pancreas Consortium

Unpublished (2000)

Other ESTs: ib44f06.x1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Dr. Hiroshi Inoue

(hinoue@wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 447.

Location/Qualifiers

Source

1. 547

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HR85 Islet"

/tissue_type="Purified pancreatic islet"

/lab_host="DH10B"

/note=Organ: Pancreas; Vector: pBluescript SK(-); Site_1:

NotI; Site_2: XhoI; cDNA made by oligo-dr priming.

Size-selected on agarose gel. Average insert size ~1kb. 5'

XhoI site was destroyed after directional cloning.

Amplified once. Contact information: Hiroshi Inoue, MD,

Metabolism Div. (Alan Permutt Lab), Washington University

School of Medicine, Box 8127, 660 South Euclid Ave., St.

Louis, MO 63110, E-mail: hinoue@ingate.wustl.edu, Tel:

314-362-1916, Fax: 314-747-2692."

BASE COUNT 152 a 120 c 174 g 101 t

ORIGIN

alignment_scores:

Quality: 653.00 Length: 125

Ratio: 5.268 Gaps: 0

Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BG655001

Align seg 1/1 to: BG655001 from: 1 to: 547

http://image.llnl.gov
 Plate: LLAM10003 row: d column: 14
 High quality sequence stop: 658.
 Location/Qualifiers

FEATURES

source
 1. .658
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4361053"
 /clone_lib="NIH_MGC_84"
 /tissue_type="adrenal cortex carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
 Note: Site_2: SalI; Cloned unidirectionally; oligo-dT
 primed. Average insert size 1.229 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 205 a 124 c 164 g 165 t
 ORIGIN

alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BF971303

Align seg 1/1 to: BF971303 from: 1 to: 658

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||||
 24 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAAATGC 73
 |||||||
 17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34
 |||||||
 74 CAACCAAGAAATCAAGAAAGAGCAAGTGTCTAATAAAGGGAGCCCT 123
 |||||||
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 |||||||
 124 TGGCCCTCCCTTTGGATGCTGTGTATATCTGTGCTAGAGAAATCGT 173
 |||||||
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIle 67
 |||||||
 174 AGGCGGTTCGCGTTAGCGAGCCATCCTGCAGTATAGATGGGATATGAT 223
 |||||||
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGlu 84
 |||||||
 224 GCATAGGCTTGGAGAACACAGCAGGAGGATGAGAGAGAGATATGGAA 273
 |||||||
 84 rGileGlyGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 |||||||
 274 GGATGGGGAGAGGTGAGACACTGATGAAAGCTGAGGAAAGAGCAG 323
 |||||||
 101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
 |||||||
 324 TTGAGTCATAGCTGCGGGCAGTCAGCACTGACCCCTCACCATGACCA 373
 |||||||
 117 sHisAspGluPheCysLeuMetPro 125
 |||||||
 374 TCATGATGAGTTTTCCTTATGCC 398
 |||||||

seq_name: gb_est2:BG714823

seq_documentation_block:

LOCUS BG714823 674 bp mRNA EST 08-MAY-2001
 DEFINITION 602677146F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4799747 5',
 mRNA sequence.
 ACCESSION BG714823
 VERSION BG714823.1 GI:13993754
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 674)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10689 row: k column: 12
 High quality sequence stop: 672.

FEATURES

source

1. .674
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4799747"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-xhoI (gtcga
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to 10⁶ 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 175 a 156 c 201 g 142 t
 ORIGIN

alignment_scores:

Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BG714823

Align seg 1/1 to: BG714823 from: 1 to: 674

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||||
 173 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAAATGC 222
 |||||||
 17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34
 |||||||
 223 CAACCAAGAAATCAAGAAAGAGCAAGTGTCTAATAAAGGGAGCCCT 272
 |||||||
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 |||||||
 273 TGGCCCTCCCTTTGGATGCTGTGTATATCTGTGCTAGAGAAATCGT 322
 |||||||
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIle 67
 |||||||
 323 AGGCGGTTCGCGTTAGCGAGCCATCCTGCAGTATAGATGGGATATGAT 372
 |||||||
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGlu 84
 |||||||
 373 GCATAGGCTTGGAGAACACAGCAGGAGGATGAGAGAGAGATATGGAA 422
 |||||||
 84 rGileGlyGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 |||||||
 423 GGATGGGGAGAGGTGAGACACTGATGAAAGCTGAGGAAAGAGCAG 472
 |||||||

101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHI 117
 473 TTGAGTCATAGTCTGGGGCAGTCAGCACTGACCCCTCACCATGACCA 522
 117 sHisAspGluPheCysLeuMetPro 125
 523 TCATGATGAGTTTGGCTTATGCC 547

seq_name: gb_est1:AV702285

seq_documentation_block:
 LOCUS AV702285 690 bp mRNA EST 08-OCT-2000
 DEFINITION AV702285 ADB Homo sapiens cDNA clone ADBXE07 5', mRNA sequence.
 ACCESSION AV702285
 VERSION AV702285.1 GI:10718615
 EST.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 690)
 AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
 TITLE Homo sapiens cDNA ADB clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zequang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 source
 1..690
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADBBXE07"
 /clone_lib="ADB"
 /tissue_type="Adrenal gland"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 210 a 132 c 180 g 166 t 2 others
 ORIGIN
 alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400
 alignment_block:
 US-09-327-750D-32 x AV702285
 Align seg 1/1 to: AV702285 from: 1 to: 690

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 60 ATGGAGTCCAAAGAAAGAACGAGGAGTAACAGTCTCAGCATGGAAATGC 109
 17 aaSngGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34
 110 CAACCAAGAAATCAAGAAAGAGCAAGTGTCTAATAAGGGGAGCCCT 159
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 160 TGGGCCCTCCCTTGGATGCTGGTGAATCTGTGCTTACAGGAATCGT 209
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67

210 AGGCGGTTCCGCTTAGGACGCCATCCCTGCAGTATAGTAGGATGAT 259
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
 260 GCATAGGCTTGGAGAACCCACAGCAAGGATGAGAGAAGAATATGGAAA 309
 84 rGileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 310 GGATTTGGGGAGGAGGTGAGACACTGATGAAAGCTGAGGAAAGCAG 359
 101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHI 117
 360 TTGAGTCATAGTCTGGGGCAGTCAGCACTGACCCCTCACCATGACCA 409
 117 sHisAspGluPheCysLeuMetPro 125
 410 TCATGATGAGTTTGGCTTATGCC 434

seq_name: gb_est1:AV705808

seq_documentation_block:
 LOCUS AV705808 691 bp mRNA EST 09-OCT-2000
 DEFINITION AV705808 ADB Homo sapiens cDNA clone ADBCRH03 5', mRNA sequence.
 ACCESSION AV705808
 VERSION AV705808.1 GI:10723103
 EST.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 691)
 AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
 TITLE Homo sapiens cDNA ADB clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zequang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 source
 1..691
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADBCRH03"
 /clone_lib="ADB"
 /tissue_type="Adrenal gland"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 211 a 133 c 181 g 162 t 4 others
 ORIGIN

alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400
 alignment_block:
 US-09-327-750D-32 x AV705808
 Align seg 1/1 to: AV705808 from: 1 to: 691

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||

```

60 ATGGAGTCCAAAGAGAAACAGCAGTAAACAGTCTCAGCATGGAAATGC 109
17 aasnGlncluasngluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
110 CAACCAAGAAATGAAGAAAGAGCAAGTGTCTAATAAAGGGAGCCCT 159
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
160 TGGCCCTCCCTTGGATGCTGTGTAATCTGTGTGCTAGAGAAATCGT 209
|||||
51 ArgArgPheProValArgGlnProLeuLeuGlnTyrArgTyrAspIleMe 67
|||||
210 AGGCGGTTCCTGGAGAGCCATCCCTGCAGTATAGATGGGATATGAT 259
|||||
67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
260 GCATAGGCTTGGAGAACACAGCAAGGATGAGAGAGAGAAATATGGAAA 309
|||||
84 rgleGlyGluGluValArgGluLeuMetGluLysLeuArgGluLysGln 100
|||||
310 GGATTTGGGAGGAGGTGAGACACTGATGGAAGAAAGCTGAGGAAAGAG 359
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
360 TTGAGTCATAGTCTGGCGGAGTGCAGCACTGACCCCTCACCATGACCA 409
|||||
117 shiAspGluPheCysLeuMetPro 125
|||||
410 TCATGATGAGTGTTCCTTATGCC 434

```

seq_name: gb_est1:AL526209

seq_documentation_block: 695 bp mRNA EST 13-FEB-2001
LOCUS AL526209 LTI_NFL003_NBC3 Homo sapiens cDNA clone CSDDC016Y001 3
DEFINITION prime, mRNA sequence.

ACCESSION AL526209
VERSION AL526209.1 GI:12789702
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES

source

```

1. .695
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDC016Y001"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: fliang@lifetech.com URL
http://fulllength.invitrogen.com"

```

BASE COUNT 158 a 198 c 154 g 183 t 2 others

ORIGIN

```

alignment_scores:
Quality: 653.00 Length: 125
Ratio: 5.266 Gaps: 0
Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:
US-09-327-750D-32 x AL526209/rev ..
Align seg 1/1 to reverse of: AL526209 from: 1 to: 695

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
546 ATGGAGTCCAAAGAGAAACAGCAGTAAACAGTCTCAGCATGGAAATGC 497
|||||
17 aasnGlncluasngluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
496 CAACCAAGAAATGAAGAAAGAGCAAGTGTCTAATAAAGGGAGCCCT 447
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
446 TGGCCCTCCCTTGGATGCTGTGTAATCTGTGTGCTAGAGAAATCGT 397
|||||
51 ArgArgPheProValArgGlnProLeuLeuGlnTyrArgTyrAspIleMe 67
|||||
396 AGGCGGTTCCTGGAGAGCCATCCCTGCAGTATAGATGGGATATGAT 347
|||||
67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
346 GCATAGGCTTGGAGAACACAGCAAGGATGAGAGAGAGAAATATGGAAA 297
|||||
84 rgleGlyGluGluValArgGluLeuMetGluLysLeuArgGluLysGln 100
|||||
296 GGATTTGGGAGGAGGTGAGACACTGATGGAAGAAAGCTGAGGAAAGAG 247
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
246 TTGAGTCATAGTCTGGCGGAGTGCAGCACTGACCCCTCACCATGACCA 197
|||||
117 shiAspGluPheCysLeuMetPro 125
|||||
196 TCATGATGAGTGTTCCTTATGCC 172

```

seq_name: gb_est2:BG820179

seq_documentation_block: 699 bp mRNA EST 22-MAY-2001

LOCUS BG820179 699 bp mRNA EST 22-MAY-2001
DEFINITION 602782222F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4933012
5', mRNA sequence.

ACCESSION BG820179

VERSION BG820179.1 GI:14167766

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 699)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: David N. Louis, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LLAM10859 row: 1 column: 05

High quality sequence stop: 695.

FEATURES

Location/Qualifiers

1. .699

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4933012"
 /clone_lib="NCI_CGAP_Brn67"
 /tissue_type="anaplastic oligodendroglioma with lp/19q loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pCMV-SpOrt6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 198 a 138 c 191 g 172 t
 ORIGIN

alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BG820179

Align seg 1/1 to: BG820179 from: 1 to: 699

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
94 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCAGTGGAAATGC 143
|||||
17 aasnGlnGluAsnGluGluGlnValAlaAsnLysGlyGluProL 34
|||||
144 CAACCAAGAAATGAAGAAAGGAGCAAGTTCCTAATAAGGGAGCCCT 193
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
194 TGGCCCTCCCTTTGGATGCTGGTGAATCTGTGCTCCTAGAGGAATCGT 243
|||||
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIle 67
|||||
244 AGCGGGTTCGGCTTAGGCAGCCATCTCTGCAGTATAGATGGATATGAT 293
|||||
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
294 GCATAGGCTTGGAGAACCAAGCAAGGATGAGAGAGAGATATGGAAA 343
|||||
84 rgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
344 GGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGGGAAGAGCAG 393
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
394 TTGAGTCATAGCTTCGGGGGAGTGCAGCATGTACCCCTCACCATGACCA 443
|||||
117 sHisAspGluPheCysLeuMetPro 125
|||||
444 TCATGATGAGTTTGGCTTATGCC 468
|||||

```

seq_name: gb_estl:BE790774

seq_documentation_block: 783 bp mRNA EST 20-SEP-2000
 LOCUS BE790774
 DEFINITION 601581956F1 NIH_MGC-7 Homo sapiens cDNA clone IMAGE:3936613 5', mRNA sequence.

ACCESSION BE790774

VERSION BE790774.1 GI:10211972

KEYWORDS EST

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 783)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue procurement: DCTD/BTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM780 row: 0 column: 14
 High quality sequence start: 3
 High quality sequence stop: 780.

FEATURES

Location/Qualifiers

1..783

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3936613"
 /clone_lib="NIH_MGC-7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"

/lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 214 a 169 c 219 g 181 t
 ORIGIN

alignment_scores:

Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BE790774

Align seg 1/1 to: BE790774 from: 1 to: 783

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
170 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGGAAATGC 219
|||||
17 aasnGlnGluAsnGluGluGlnValAlaAsnLysGlyGluProL 34
|||||
220 CAACCAAGAAATGAAGAAAGGAGCAAGTTCCTAATAAGGGAGCCCT 269
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
270 TGGCCCTCCCTTTGGATGCTGGTGAATCTGTGCTCCTAGAGGAATCGT 319
|||||
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIle 67
|||||
320 AGCGGGTTCGGCTTAGGCAGCCATCTCTGCAGTATAGATGGATATGAT 369
|||||
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
370 GCATAGGCTTGGAGAACCAAGCAAGGATGAGAGAGAGATATGGAAA 419
|||||
84 rgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
420 GGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGGGAAGAGCAG 469
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
470 TTGAGTCATAGCTTCGGGGGAGTGCAGCATGTACCCCTCACCATGACCA 519
|||||
117 sHisAspGluPheCysLeuMetPro 125
|||||

```

520 TCATGATGAGTTTTCCTTATGCC 544

seq_name: gb_est1:AL526247

seq_documentation_block: 794 bp mRNA EST 13-FEB-2001
LOCUS AL526247 LTI_NFL003.NBC3 Homo sapiens cDNA clone CS0DC016Y001 5
DEFINITION prime, mRNA sequence.

ACCESSION AL526247
VERSION AL526247.1 GI:12789740
KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 794)

AUTHORS Li, W. B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

1..794

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="LTI_NFL003_NBC3"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(GT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 236 a 163 c 213 g 180 t

ORIGIN

alignment_scores:

Quality: 653.00 Length: 125

Ratio: 5.266 Gaps: 0

Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x AL526247

Align seg 1/1 to: AL526247 from: 1 to: 794

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

148 ATGGAGTCCAAAGAGAACGAGCAGTAAACAGCTCAGCATGGAAATGC 197

17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34

198 CAACCAAGAAATGAAGAAGAGGAGCAAGTTGCTAATAAAGGGAGCCCT 247

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50

248 TGGCCCTCCCTTGGATGCTGGTGAATACCTGTGCTAGAGGAATCGT 297

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTyrAspIleWe 67

298 AGCGCGTTCCCGTTCAGGAGCCCATCTGCAGTATAGATGGATATGAT 347

67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84

348 GCATAGGCTTGAGAACACAGGAGGATGAGAGAGATATGAAA 397

84 rgIleGlyGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100

398 GGATTGGGAGAGGCTGAGACAGCTGATGGAAAAAGCTGAGGAAAAGCAG 447

101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspH 117

448 TTGAGTCATAGTCTCGGGGAGTCTGAGCTGACCCCTCACCATGACCA 497

117 shisAspGluPheCysLeuMetPro 125

498 TCATGATGAGTTTTCCTTATGCC 522

seq_name: gb_est1:AV702643

seq_documentation_block: 683 bp mRNA EST 08-OCT-2000

LOCUS AV702643 ADB Homo sapiens cDNA clone ADBDBE11 5', mRNA sequence.

DEFINITION AV702643 ADB Homo sapiens cDNA clone ADBDBE11 5', mRNA sequence.

ACCESSION AV702643

VERSION AV702643.1 GI:10718973

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 683)

AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao,

H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,

Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu,

G., Hu, R., Chen, J., Chen, Z., and Han, Z.

TITLE Homo sapiens cDNA ADB clones

JOURNAL Unpublished (2000)

COMMENT Contact: Zengqiang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzq@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

Location/Qualifiers

1..683

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="ADBDBE11"

/tissue_type="Adrenal gland"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 208 a 128 c 184 g 161 t

ORIGIN

alignment_scores:

Quality: 650.00 Length: 125

Ratio: 5.242 Gaps: 0

Percent Similarity: 99.200 Percent Identity: 97.600

alignment_block:

US-09-327-750D-32 x AV702643

Align seg 1/1 to: AV702643 from: 1 to: 683

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

60 ATGGAGTCCAAAGAGAACGAGCAGTAAACAGCTCAGCATGGAAATGC 109

17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34

Email: cgaabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10679 row: p column: 05
High quality sequence stop: 712.
High quality sequence stop: 712.

FEATURES
source
1. 734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="479582"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R05. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 187 a 174 c 214 g 159 t
ORIGIN

alignment_scores:
Quality: 646.00 Length: 125
Ratio: 5.252 Gaps: 0
Percent Similarity: 98.400 Percent Identity: 97.600

alignment_block:
US-09-327-750D-32 x BG715659

Align seg 1/1 to: BG715659 from: 1 to: 734

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
206 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGGAATGCG 255
17 aasnGlnGluAsnGluLysGluLysGluValAlaAsnLysGlyGluProL 34
|||||
256 CACCAAGAAATGAGGAAGGAGGAGTCTGCTATTAAGGGAGGCCCT 305
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
306 TGCCCTCCCTTTGGATGCTGCTGAATGCTGCTGCTAGAGGAAATCGT 355
51 ArgArgPheProValArgGlnProLeuGlnTyrArgTyrPaspIleW 67
|||||
356 AGCGGGTCCGGCTGAGCAGCCCATCTCGCATATAGATGGATATGAT 405
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
406 GCATAGGCTTGGAGAACCCACAGGAGGATGAGAGAGAAATATGAAA 455
84 rgtleGlyGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
456 GGNATGGGAGGAGGTGAGACAGCTGATGGAAGACCTGAGGGAAGACG 505
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHI 117
|||||
506 TTGAGTCATAGTCTCGGCGCAGTACGACTGACCCCTCACCATGACCA 555
117 sHisAspGluPheCysLeuMetPro 125
|||||
556 TCATGATGAGTTTTCCTTATGCCC 580

seq_name: gb_est2:BG709325
seq_documentation_block:
LOCUS BG709325 828 bp mRNA EST 07-MAY-2001
DEFINITION 602673345F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:479582 5',
mRNA sequence.
ACCESSION BG709325
VERSION BG709325.1 GI:13987550
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10679 row: j column: 11
High quality sequence stop: 798.
High quality sequence stop: 798.

FEATURES
source
1. 828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="479582"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R05. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 232 a 179 c 228 g 189 t
ORIGIN

alignment_scores:
Quality: 637.00 Length: 126
Ratio: 5.137 Gaps: 1
Percent Similarity: 98.413 Percent Identity: 96.825

alignment_block:
US-09-327-750D-32 x BG709325

Align seg 1/1 to: BG709325 from: 1 to: 828

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
176 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGGAATGCG 225
17 aasnGlnGluAsnGluLysGluLysGluValAlaAsnLysGlyGluProL 34
|||||
226 CACCAAGAAATGAGGAAGGAGGAGTCTGCTATTAAGGGAGGCCCT 275
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
276 TGCCCTCCCTTTGGATGCTGCTGAATGCTGCTGCTAGAGGAAATCGT 325


```

; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGILIO, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-78

alignment_scores:
    Quality: 84.00      Length: 102
    Ratio: 1.355        Gaps: 4
    Percent Similarity: 60.784      Percent Identity: 32.353

alignment_block:
US-09-327-750D-32 x PCT-US93-06251-78
Align seg 1/1 to: PCT-US93-06251-78 from: 1 to: 2301

2 GluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAs 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 GAGGAGCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 126

18 nGInGluAsnGluGluLysGluInValAlaAsnLysGlyGluProLeuA 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTG 176

35 laLeuProLeuAspIaGlyGluTyrCysValProArgGlyAsnArgArg 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:
177 CAACGCCGAG.....CGGAGGAGG 195

52 ArgPheProValArgGlnProIleLeu.....GlnTyrArgTrpAspI 66
::: ||| |||:|||||:|||||:|||||:|||||:|||||:
196 CAGGAACCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 245

66 eMethHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMet 82
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 CCACCATCGAGCGGTGAAGCGCAGATCCAGGTCTGCAGCAGCAGGCA 295

83 GluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluL 99
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
296 GATGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 345

99 ys 99
||
346 AA 347

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-872-644-52
seq_documentation_block:
; Sequence 52, Application US/07872644
; Patent No. 5389527
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

```

```

; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,644
; FILING DATE: 19920420
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5389527and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
US-07-872-644-52

alignment_scores:
    Quality: 80.50      Length: 126
    Ratio: 1.118        Gaps: 6
    Percent Similarity: 57.143      Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-07-872-644-52
Align seg 1/1 to: US-07-872-644-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnGl 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
959 CAGAGGCGTTCGAGTTTGAATAGCATCAGCTCGTCAGATGCCAAGCGATC 1008

20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
|||||:|||||:|||||:|||||:|||||:|||||:
1009 AGGTGTCAAGACCTCTGGTTTCAGAGGAGTCCCGGATCAACAATCTG 1058

31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1059 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTTACTTTGGACGGAAGTGTG 1108

44 sval.....ProArgLysAsnArgArgPheP 54
|||||:|||||:|||||:|||||:|||||:|||||:
1109 CACATCAATCGGAGAGATGGAGGCGCCAGGTACCCTCAAGAGGAGGAGGC 1158

54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....IleMethHis 68
|||:|||||:|||||:|||||:|||||:|||||:|||||:
1159 CAAGAAGGAAGCAGAGAGGAAAAGGCTCGCTGCGCCGACAGAGGAGCAAA 1208

69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgIl 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1209 AGGAATGGAAGCCAAAGCCAGGCTGAAGAGGCGCATCTGGCAAGCT 1258

```

85 eGlyGluValuArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
 1259 GAGA.....AAAGAGCTGGGAGAACTAAGATCAAGTCAATGG 1299
 102 erHisSerLeuArgAlaValSerThr 110
 1300 AACACGGGCAACAAAGTGACAACC 1325

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-494-52

```
seq_documentation_block:
; Sequence 52, Application US/08297494
; Patent No. 5580771
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,494
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5580771and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
US-08-297-494-52
```

alignment_scores:
 Quality: 80.50 Length: 126
 Ratio: 1.118 Gaps: 6
 Percent Similarity: 57.143 Percent Identity: 24.603

alignment_block:
 US-09-327-750D-32 x US-08-297-494-52

Align seg 1/1 to: US-08-297-494-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG 20
 959 CAGAGCGGTTCAGTTTGAATAGCATCAGCTCGTCAGATGCCAAGGATC 1008
 20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
 1009 AGGTGTCAAGACCTCTGTTTCAGAGGGAAGTCCCGCATCAACAATCTG 1058
 31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrcy 44
 1059 TCATCTCGTTGACTATAGAGCTTTAAAGCTACTTGGACGGAGTGGTG 1108
 44 sval.....ProArgGlyAsnArgArgPhep 54
 1109 CACATCAATCGGGAGAGATGAGGGCCAAAGGTACCCAAAGAGGAGGCGC 1158
 54 rovalArgGlnProIleLeuGlnTyArgTrpasp.....IleMetHis 68
 1159 CAAGAAGGAAGCAGAGAGGAAAGGCTCGCTGGCCGACAGAGCAGCACA 1208
 69 ArgLeuGlyGluProGlnAlaArgMetArgGluGlnAsnMetGluArg 85
 1209 AGGAAATGGAAGCCCAAAAGCCAGGCTGAAGAAGCGGCATCTGCAAGCT 1258
 85 eGlyGluGluValuArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
 1259 GAGA.....AAAGAGCTCTGGAGAACTAAGATCAAGTCAATGG 1299
 102 erHisSerLeuArgAlaValSerThr 110
 1300 AACACGGGCAACAAAGTGACAACC 1325

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-510-52

```
seq_documentation_block:
; Sequence 52, Application US/08297510
; Patent No. 5602019
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,510
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5602019and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 346-5750
```

```

; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
; US-08-297-510-52

alignment_scores:
    Quality: 80.50      Length: 126
    Ratio: 1.118        Gaps: 6
    Percent Similarity: 57.143    Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-08-297-510-52

Align seg 1/1 to: US-08-297-510-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
|||||
959 CAGAGCGCTTCGAGTTGATAGATCATCAGCTCGTCAGATGCCAAGCGATC 1008
20 u.....AsnGluGluLysGluInValAlaAsnLys. 30
1009 AGGTGTCAGACCTCTCGTTCAGAGGGAAGTCCCGCATCAACAATTCG 1058
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
|||||
1059 TCATCTCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1108
44 sVal.....ProArgGlyAsnArgArgArgPheP 54
|||||
1109 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGGGAAGGC 1158
54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....IleMethis 68
|||||
1159 CAAGAAGCAAGCAGAGAAAGGCTCGCTGGCGCAGAGGAGCAGCAAA 1208
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgil 85
|||||
1209 AGGAATGGAAGCCAAAGCCAGCTGAAGAAGCGCATCTGGCAAAGCT 1258
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeus 102
|||||
1259 GAGA.....AAAAGCTCTGGAGAACTAAGAATCAAGTCAATGG 1299
102 erHisSerLeuArgAlaValSerThr 110
|||||
1300 AACACGGGCAACAAAAGTGACAAACC 1325

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-479-532-52

seq_documentation_block:
; Sequence 52, Application US/08479532
; Patent No. 5776752
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF INVENTIONS: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell

```

```

; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,532
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5776752and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
; US-08-479-532-52

alignment_scores:
    Quality: 80.50      Length: 126
    Ratio: 1.118        Gaps: 6
    Percent Similarity: 57.143    Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-08-479-532-52

Align seg 1/1 to: US-08-479-532-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
|||||
959 CAGAGCGCTTCGAGTTGATAGATCATCAGCTCGTCAGATGCCAAGCGATC 1008
20 u.....AsnGluGluLysGluInValAlaAsnLys. 30
1009 AGGTGTCAGACCTCTCGTTCAGAGGGAAGTCCCGCATCAACAATTCG 1058
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
|||||
1059 TCATCTCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1108
44 sVal.....ProArgGlyAsnArgArgArgPheP 54
|||||
1109 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGGGAAGGC 1158
54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....IleMethis 68
|||||
1159 CAAGAAGCAAGCAGAGAAAGGCTCGCTGGCGCAGAGGAGCAGCAAA 1208
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgil 85
|||||

```

```

US-09-327-750D-32 x US-08-455-526-52      ..
Align seg 1/1 to: US-08-455-526-52 from: 1 to: 2077

  4  LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  959 CAGAGGCGTTCAGGTTGGAATAGCATCAGCTCGTCAGATGCCAAGCGCATC 1008
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1009 AGGTGTCACAGACCTCTGGTTCAGAGGGGAAGTGCCTCCCATCAACAATTCG 1058
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1059 TCATCTCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      44 sVal.....ProArgLysAsnArgArgArgPheP 54
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1109 CACATCAATCGGGAGAGATGGAGGCGCAAGGTACCCAAAGAGGAGGAAGC 1158
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      54 roValArgGlnProIleLeuGlnTyrArgTTrpAsp.....IleMethHis 68
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1159 CAGAAGGAAGCAGAGAAAGGCTCGCCTGCGCCGAGAGGAGCAGCAAA 1208
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI1 85
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1209 AGGAATGGAGGCCAAAAGCCAGGCTGAAGAAGCGCGCATCTGCCAAGCT 1258
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeu 102
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1259 GAGA.....AAAGACGCTCTGGAGAAACTAAGAATCAAGTCAATATGG 1299
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  102 eHisSerLeuArgAlaValSerThr 110
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  1300 AACACGGGCAAAACAAAGTGACAACC 1325
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

seq name: /cqn2 6/ptodata/2/ina/5A_COMB_seq:US-08-455-525-52

```

```

seq_documentation_block:
; Sequence 52, Application US/08455525
; Patent No. 5800987
;
; GENERAL INFORMATION:
;
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
;
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
;
; NUMBER OF SEQUENCE: 58
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,525
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:

```

```
alignment_scores:
  Quality: 80.50
  Ratio: 1.118
  Percent Similarity: 57.143
alignment_block:
  Length: 126
  Gaps: 6
  Percent Identity: 24.603
```

```

; NAME: No. 5800987and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
; US-08-455-525-52

alignment_scores:
    Quality: 80.50      Length: 126
    Ratio: 1.118       Gaps: 6
    Percent Similarity: 57.143      Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-08-455-525-52
..
Align seg 1/1 to: US-08-455-525-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
959 CAGAGCGGTTGAGTTGAATAGCATCAGCTCGTCAGATGCCAAGCGATC 1008
20 u.....AsnGluLysGluGlnValAlaAsnLys. 30
1009 AGGTGTCAGACCTCTGGTTTCAGAGGAGGAGTCCCCCGATCAACAATCTG 1058
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
1059 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1108
44 sVal.....ProArgGlyAsnArgArgArgPhep 54
1109 CACATCAATCGGAGAGATGAGGCGCAAGTACCCCAAGAGAGAGAGGC 1158
54 roValArgGlnProLeuGlnTyrArgTrpasp.....IleMetHis 68
1159 CAGAAGGAGGAGAGAGAGGAGGAGGAGTCCGCTGCGCGCAGAGAGAGCAAA 1208
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgil 85
1209 AGGAATGGAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1258
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluGluGlnLeus 102
1259 GAGA.....AAAGACCTCTGGAGAACTAAGAACTCAAGTCAATGG 1299
102 erHisSerLeuArgAlaValSerThr 110
1300 AACAGCGGCAACAAACAAAGTGACAAAC 1325

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-139-491-52

seq_documentation_block:
; Sequence 52, Application US/09139491
; Patent No. 6015677
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian

```

54 roValArgGlnProIleLeuGlnTyrArgTirpasp.....lleMetHis 68
|||
1159 CAAGAAGGAGCAGAGGAAAAGGCTCCCTGGCGCAGAGGAGGAGCAAA 1208
|||
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI 85
|||
1209 AGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1258
|||
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeu 102
|||
1259 GAGA.....AAAAGAGCTCTGGAGAACTAAGAATCAAGTCAATGG 1299
|||
102 erHisSerLeuArgAlaValSerThr 110
|||
1300 AACAGGGGCAACAAAGTGCACACC 1325

seq_name: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq: PCT-US92-03222-52

seq_documentation_block:
; Sequence 52, Application PC/TUS9203222
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/03222
; FILING DATE: 19920420
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
PCT-US92-03222-52

alignment_scores:
Quality: 80.50 Length: 126

Ratio: 1.118 Gaps: 6
Percent Similarity: 57.143 Percent Identity: 24.603
alignment_block:
US-09-327-750D-32 x PCT-US92-03222-52 ..
Align seg 1/1 to: PCT-US92-03222-52 from: 1 to: 2077
4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG 20
|||
959 CAGAGCGCTTCAGTTTGATAGCATCAGCTCGTCGATGCCAAGCGATC 1008
|||
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
|||
1009 AGGTGTCAAGACCTCTGTTTCAGAGGAGGAGTCCCGCATCAACAATT 1058
|||
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
|||
1059 TCATCTCGTTGACTATAAGAGCTTTAAAGCTACTTTGGACGGAGTGG 1108
|||
44 sVal.....ProArgGlyAsnArgArgArgPhep 54
|||
1109 CACATCAATCGGAGAGATGGAGGCGCAAGGTACCCAAAGAGAGAGAGC 1158
|||
54 roValArgGlnProIleLeuGlnTyrArgTirpasp.....lleMetHis 68
|||
1159 CAAGAAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 1208
|||
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArg 85
|||
1209 AGGAATGGAAGCCAAAGCCAGGCTGAGAGAGGCGCATCTGGCAAGCT 1258
|||
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeu 102
|||
1259 GAGA.....AAAAGAGCTCTGGAGAACTAAGAATCAAGTCAATGG 1299
|||
102 erHisSerLeuArgAlaValSerThr 110
|||
1300 AACAGGGGCAACAAAGTGCACACC 1325

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq: US-07-872-644-50

seq_documentation_block:
; Sequence 50, Application US/07872644
; Patent No. 5389527
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,644
; FILING DATE: 19920420
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/688,356

```

FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5389527and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2693 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 176..2077
US-07-872-644-50

alignment_scores:
Quality: 80.50 Length: 126
Ratio: 1.118 Gaps: 6
Percent Similarity: 57.143 Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-07-872-644-50 ..
Align seg 1/1 to: US-07-872-644-50 from: 1 to: 2693

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
1568 CAGAGCGCTTCGAGTTGAATAGCATCAGCTCGTCAGATGCCAAGCGATC 1617
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
1618 AGGTCTCAAGACCTCTGTTTCAGAGGGAAGTCCCGCATCAACATCTG 1667
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
1668 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1717
44 sval.....ProArgGlyAsnArgArgArgPhep 54
1718 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGGAGGAGGC 1767
54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....IleMetHis 68
1768 CAAGAAGGAGGAGAGGAGGAAAGGCTCGCTGGCGCAGAGGAGCAGCAAA 1817
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI1 85
1818 AGGAATGGAAGCAAAAGCCAGGCTGAAGAGCGGCATCTGCCAAGCT 1867
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
1868 GAGA.....AAAGACGCTGGAGAAACTAAGAATCAAGTCAATGG 1908
102 eHisSerLeuArgAlaValSerThr 110
1909 AACACGGGCAACAAAGTGACACC 1934

seq_name: /cgn2/6/ptodata/2/1na/5A_COMB.seq:US-08-297-494-50
seq_documentation_block:
Sequence 50: Application US/08297494
Patent No. 5580771
GENERAL INFORMATION:
APPLICANT: Beavo, Joseph A.
APPLICANT: Bentley, Kelley
APPLICANT: Charbonneau, Harry

```

```

APPLICANT: Sonnenburg, William K.
TITLE OF INVENTION: DNA Encoding Mammalian
Phosphodiesterases
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/297,494
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/688,356
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 558077land, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2693 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 176..2077
US-08-297-494-50

alignment_scores:
Quality: 80.50 Length: 126
Ratio: 1.118 Gaps: 6
Percent Similarity: 57.143 Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-08-297-494-50 ..
Align seg 1/1 to: US-08-297-494-50 from: 1 to: 2693

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
1568 CAGAGCGCTTCGAGTTGAATAGCATCAGCTCGTCAGATGCCAAGCGATC 1617
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
1618 AGGTCTCAAGACCTCTGTTTCAGAGGGAAGTCCCGCATCAACATCTG 1667
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
1668 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1717
44 sval.....ProArgGlyAsnArgArgArgPhep 54
1718 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGGAGGAGGC 1767
54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....IleMetHis 68

```


[illegible]

```

1668 TCATCTCCGTTGACTATAAGACCTTTAAAGCTTACTTCGCAGGAAGTGCGT 1717
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::

    44 sval.....ProArgIyAsnArgArgPhep 54
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

1718 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGNGRAGGC 1767
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

    54 roValArgGlnProIleLeuGlnTyArgTrpasp.....IleMethIs 68
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

1768 CAGAAGAGGAAGCAGAGGAAAAGGCTCGCTCGCCGAGAGGAGCAGCAA 1817
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

    69 ArgLeuGIYGLUProgInAlaArgMetArgGluiAsnMetGIUAIGIL 85
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

1818 AGGAATGGAACCCAAAGCCAGGCTGAAGAAGCGCATCTGCCARAAGCT 1867
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

    85 eGIYGLUIGLUValArgGlnLeuMetGIULyLysLeuArgGIULyGlnLeus 102
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

1868 GAGA.....AAAAAGCCTCTGGAGAACTTAAGAAATCAAGTCAATGG 1908
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

102 exHisSerLeuArgAlaValSerThr 110
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::
      :|::|||::|||::|||::|||::|||::|||::|||::|||::

1909 AACACGGGCAAAACAAAGTGTACAACC 1934
      :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::

```

Sequence 50, Application US/08479532
Patent No. 5776752

GENERAL INFORMATION:

APPLICANT: Beavo, Joseph A.
APPLICANT: Bentley, Kelley
APPLICANT: Charbonneau, Harry
APPLICANT: Sonnenburg, William K.
TITLE OF INVENTION: DNA Encoding Mammalian
TITLE OF INVENTION: Phosphodiesterases
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,532
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/297,494
FILING DATE:

```
alignment_scores:
  Quality: 80.50
  Ratio: 1.118
Length: 126
Gaps: 6
```

Tue Mar 12 09:01:37 2002

```

APPLICATION NUMBER: US 07/688,356
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5776752and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2693 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 176..2077
US-08-479-532-50

alignment_scores:
Quality: 80.50 Length: 126
Ratio: 1.118 Gaps: 6
Percent Similarity: 57.143 Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-08-479-532-50

Align seg 1/1 to: US-08-479-532-50 from: 1 to: 2693

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
1568 CAGAGCGTTCGAGTTTCAATAGCATCAGCTCGTCAGATGCCAAGCGATC 1617
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
1618 AGGTGTCAAGACCTCTGGTTCAGAGGGAAGTCCCCCGATCAACAATCTG 1667
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
1668 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTACTTGGACGGAAGTGGTG 1717
44 sVal.....ProArgGlyAsnArgArgPheP 54
1718 CACATCAATCGGAGAGATGGAGGCCCAAGGTACCCAAAGAGGAGAGGC 1767
54 roValArgGlnProIleLeuGlnTyrArgTyrPasp.....IleMetHis 68
1768 CAAGAAGGAAGCAGAGAGAAAGGCTCGCTGGCCGACAGAGGAGCAGCAA 1817
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArg11 85
1818 AGGAATGGAGCCAAAGCCAGGCTGAAGAGGCGGCATCTGGCAAGCT 1867
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeus 102
1868 GACA.....AAAAGACGCTCTGAGAAACTAAGAATCAAGTCAATGG 1908
102 erHisSerLeuArgAlaValSerThr 110
1909 AACACGGGCAACAAAGTGCACACC 1934
```

OM of: US-09-327-750D-32 to: N_Geneseq_1101.* out_format : pfs
 Date: Mar 11, 2002 3:42 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
 -MODEL=framet_p2n.model -DEV=xlh
 -Q/cgna2_1/USP10.spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
 -DB=N_Geneseq_1101 -QFMT=fastac -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
 -USR=US09327750 -CGEN1_1_330 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-09-327-750D-32
 Query length: 125
 Database: N_Geneseq_1101.*
 Database sequences: 930621
 Database length: 428662619
 Search time (sec): 355.560000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143		639.00	1250.82	792	8.0e-64
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810		602.50	1216.52	767	1.1e-59
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58581		602.50	1215.36	862	1.3e-59
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611		602.50	1214.95	898	1.3e-59
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH60367		589.50	1188.82	858	3.8e-58
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH30880		585.50	1183.24	792	7.9e-58
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750		585.50	1182.71	698	8.4e-58
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH30879		562.50	1135.77	692	3.5e-55
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH10889		207.50	414.96	421	4.9e-15
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06183		207.50	414.14	421	5.4e-15
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03517		202.50	397.55	865	4.6e-14
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750		202.50	394.04	1229	7.2e-14
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH85548		202.50	393.00	1364	8.2e-14
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH23529		201.00	394.18	891	7.0e-14
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH23528		185.50	362.21	917	4.2e-12
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH23528		185.50	363.88	700	3.4e-12
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06184		184.50	367.25	451	2.2e-12
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01005		151.50	298.13	532	1.6e-08
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06181		117.00	234.04	485	7.1e-07
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01003		106.00	207.93	401	5.8e-05
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH87147		100.50	204.30	187	0.0026
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH14053		89.00	160.28	1459	0.7498
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82586		88.50	171.55	426	0.1767
/SID2/gcgdata/geneseq/geneseq/NA1992.DAT:AAH28398		86.50	141.15	5946	8.73
/SID2/gcgdata/geneseq/geneseq/NA1992.DAT:AAH70447		84.00	134.79	11236	19.71
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH70445		84.00	145.51	2301	4.99
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH09303		81.00	135.10	3532	4.99
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15730		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1992.DAT:AAH03188		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH083982		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH67203		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH51117		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH47866		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH48209		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH36161		80.50	139.38	2077	10.95
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH290385		80.50	136.79	2693	15.27
/SID2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH083981		80.50	136.79	2693	15.27
/SID2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH67202		80.50	136.79	2693	15.27
/SID2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH51116		80.50	136.79	2693	15.27

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143
 seq_documentation_block:
 ID AAH45143 standard; cDNA; 792 BP.
 AC AAH45143;
 XX
 DT 07-SEP-2001 (first entry)
 DE Human brain expressed X-linked protein, hbex, coding sequence.
 XX
 DE Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;
 KW hbex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;
 KW X-chromosome-binding mental retardation; lissencephalous disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200140286-A1.
 XX
 PD 07-JUN-2001.
 XX
 PF 27-NOV-2000; 2000WO-CN00502.
 XX
 PR 30-NOV-1999; 99CN-0124179.
 XX
 PA (BIOR) BIOROAD GENE DEV LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 DR WPI; 2001-397944/42.
 DR P-PSDB; AAB99224.
 XX
 PT Isolated human brain-expressed X-linked polypeptide used to diagnose
 and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
 PT deafness and X-chromosome-binding mental retardation
 XX
 PS Claim 5; Page 22; 30pp; Chinese.
 XX
 CC The present sequence is the coding sequence for a human brain-expressed
 X-linked protein (hbex). hbex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, X-chromosome-binding mental retardation and
 CC lissencephalous disease. hbex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hbex coding sequence can be used as primers or probes, or in producing
 gene chips or microarrays.
 XX
 SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;
 alignment_scores:
 Quality: 639.00 Length: 125
 Ratio: 5.238 Gaps: 0
 Percent Similarity: 97.600 Percent Identity: 96.800
 alignment_block:
 US-09-327-750D-32 x AAH45143
 Align seg 1/1 to: AAH45143 from: 1 to: 792
 1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 173 ATGAGTCCAAAGAGAACTAGCAGTAAACAGTCTCAGCATGGAATGC 222
 17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
 223 CAACCAAGAAATGAGAGAAAGGAGCAAGTCTGCTAATAAAGGGAGCCCT 272

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 |||||
 273 TGGCCCTCCCTTTGGATGCTGGTGAATACATCTGCTAGAGGAATCGT 322
 |||||
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTyrAspIleMe 67
 |||||
 323 AGGCGGTTCGCGGTAGGACGCCCATCTGCAGTATAGATGATGAT 372
 |||||
 67 HisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
 |||||
 373 GCATAGGCTTGAGAACACACAGGAGGATGAGAGAGAGATATGGAAA 422
 |||||
 84 TgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 |||||
 423 GCATTTGGGAGGGGTGAGACAGCTGATGAAAAGCTGAGGAAAAGCAG 472
 |||||
 101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspH 117
 |||||
 473 TTGAGTCACTCTGCGGGCAGTCAGCACTGACCCCTCACCATGACCA 522
 |||||

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH75810

seq_documentation_block:
 ID AAH75810 standard; cdna; 767 BP.

XX AC AAH75810;

XX DT 17-OCT-2001 (first entry)

XX DE Human X chromosome linked gene expression protein 14 coding sequence.

XX KW Human; X chromosome linked gene expression protein 14; cancer;

XX KW HIV infection; cytosolic; anti-HIV; chromosome X; ss.

XX OS Homo sapiens.

XX PN CN1296969-A.

XX PD 30-MAY-2001.

XX PF 23-NOV-1999; 99CN-0124078.

XX PR 23-NOV-1999; 99CN-0124078.

XX PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX PI Mao Y, Xie Y;

XX DR WPI: 2001-483897/53.

XX DR P-PSDB: AAG66407.

XX PT Polypeptide-human X chromosome linked gene expression protein 14 and polynucleotide for coding said polypeptide -

XX PS Claim 6; Page 24 (Disclosure); 31pp; Chinese.

XX CC The present sequence is the coding sequence for human X chromosome linked gene expression protein 14. The protein and coding sequence are useful for treating diseases e.g. cancer and HIV infection.

XX SQ Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;

alignment_scores:

Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x AAH75810 ..

Align seg 1/1 to: AAH75810 from: 1 to: 767

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||
 101 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCGTGGAAAATG 150
 |||||
 17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 |||||
 151 CAACCCAGGAAATATGATGAAAAGATGAAAAGGAGCAAGTTGCTAATAAG 200
 |||||
 31 LysGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||
 201 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTAATCTGTGTGCCTAGA 250
 |||||
 48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
 |||||
 251 GGAACACCTAGCGGTTCCGCTTAGCAGCCCATCTGCAGTATAGATG 300
 |||||
 64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluAla 81
 |||||
 301 GGACATAATGCTAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 350
 |||||
 81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 |||||
 351 ATATGGAAGGATTTGGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGG 400
 |||||
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 114
 |||||
 401 GAAAAGCAGATTGAGTCATAGTCTCGGGCAGTCAGCACACTGATCCCCCTCA 450
 |||||
 114 shisAspHisHisAspGluPheCysLeuMetPro 125
 |||||
 451 CCATGACCATCATCAGTATGATTTTGCCTTATGCCCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI58581

seq_documentation_block:

ID AAI58581 standard; cdna; 862 BP.

XX AC AAI58581;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 784.

XX KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM39425.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Claim 1; SEQ ID NO 784; 10078pp; English.
 PS The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic, and
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:
 Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062
 alignment_block:
 US-09-327-750D-32 x AAI58581
 Align seg 1/1 to: AAI58581 from: 1 to: 862
 1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 226 ATGGAGTCCAAAGGAGGAGCGCTTAACAATCTCATCGTGGAATGCT 275
 17 abasnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 276 CAACAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTTAATAAG 325
 31 lYGluproLeuAlaLeuproLeuAspAlaGlyGluTyrCysValProArg 47
 326 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTACTAGTACTGTGCTAGA 375
 48 GlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
 376 GGAACCCGTAGCGGTTCCTCGGTTAGGACGCCATCTCGCAGTATAGATG 425
 64 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGlu 81
 426 GCACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 475
 81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 476 ATATGGAAGAGGATGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGG 525
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 114
 526 GAAAGGAGTGTGATCATAGTTTCGGGCGAGTCAGCACTGATCCCTCA 575
 114 shIsAspHisAspGluPheCysLeuMetPro 125
 ||||||||||||||||||||||||||||||||||||||||||||

576 CCATGACCATCAGTACGATGAGTTTTCGCTTATGCCC 609

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAF59611

seq_documentation_block:

ID_AAF59611 standard; cDNA; 898 BP.

XX AAF59611;

XX 24-APR-2001 (first entry)

XX Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.

XX Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection; ss.

XX Homo sapiens.

XX WO200107471-A2.

XX 01-FEB-2001.

XX 21-JUL-2000; 2000WO-US19948.

XX 21-JUL-1999; 99US-0145075.

XX 08-SEP-1999; 99US-0153129.

XX 10-NOV-1999; 99US-016447.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;

XX Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;

XX WPI: 2001-112727/12.

XX P-PSDB; AAB60474.

XX Human cell cycle and proliferation proteins and polynucleotides are
 PT used to treat, diagnose and prevent immune, developmental and cell
 PT signaling disorders and cell proliferative disorders including cancer -
 PS Claim 5; Page 181-182; 205pp; English.

CC Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
 CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in drug screening methods to
 CC identify compounds that modulate the activity of CCYPR. CCYPR
 CC nucleotides can be used to generate transgenic animal models of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC Diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders, and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.

XX Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

alignment_scores:
 .. Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:
 US-09-327-750D-32 x AAF59611 ..

Align seg 1/1 to: AAF59611 from: 1 to: 898

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||.....:|||||.....:|||||.....:|||||.....:
254 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCGTGGAAATGT 303
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
:|||||.....:|||||.....:|||||.....:|||||.....:
304 CAACCGAGAAATGATGAATAAGATGAAAGAGGCAAGTTGCTATAAAG 353
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||.....:|||||.....:|||||.....:|||||.....:
354 GGGAGCCCTGGCCCTACCTTGAATGTTAGTGAATCTGTGCTAGAGA 403
48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
|||||.....:|||||.....:|||||.....:|||||.....:
404 GGAACCGTAGGGGTTCCGCGTTAGGCAGCCCATCTCGAGTATAGATG 453
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluAla 81
|||||.....:|||||.....:|||||.....:|||||.....:
454 GCACATATCATAGCTTGGAGGACCCACAGCAGGATGAGAGGAGAGA 503
81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
|||||.....:|||||.....:|||||.....:|||||.....:
504 ATATGGAAGGATTGGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGG 553
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
|||||.....:|||||.....:|||||.....:|||||.....:
554 GAAAGCAGTTGAGTCATAGTCTGCGGGCAGTCAGCAGTATCCCTCA 603
114 sHisAspHisHisAspGluPheCysLeuMetPro 125
|||||.....:|||||.....:|||||.....:|||||.....:
604 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 637

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI60367

seq_documentation_block:
 ID AAI60367 standard; cDNA: 858 BP.

XX AAI60367;
 XX 22-OCT-2001 (first entry)
 XX Human polynucleotide SEQ ID NO 4356.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX Homo sapiens.
 XX W0200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US34263.
 XX 21-JAN-2000; 2000US-0498725.
 XX 25-APR-2000; 2000US-0552317.
 XX 09-JUL-2000; 2000US-0598042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM41211.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Claim 1; SEQ ID NO 4356; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAI42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 858 BP; 236 A; 175 C; 241 G; 206 T; 0 other;
 SQ

alignment_scores:
 Quality: 589.50 Length: 129
 Ratio: 4.793 Gaps: 2
 Percent Similarity: 95.349 Percent Identity: 88.372

alignment_block:
 US-09-327-750D-32 x AAI60367 ..

Align seg 1/1 to: AAI60367 from: 1 to: 858

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||.....:|||||.....:|||||.....:|||||.....:
215 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCGTGGAAATGT 264
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
:|||||.....:|||||.....:|||||.....:|||||.....:
265 CAACCGAGAAATGATGAATAAGATGAAAGAGGCAAGTTGCTATAAAG 314
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||.....:|||||.....:|||||.....:|||||.....:
315 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGCTAGAGA 364
48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
|||||.....:|||||.....:|||||.....:|||||.....:
365 GGAACCCGTAGCGGTTCCGCGTTAGGCAGCCCATCTCTCAGTATAGATG 414
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
|||||.....:|||||.....:|||||.....:|||||.....:
415 GGACATAATGATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 464
81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
|||||.....:|||||.....:|||||.....:|||||.....:
465 ATATGGAAGGATTGGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGG 514
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114

```

|||||
515 GAAAGCAGTTGAGTCATAGTCTGGGCGAGTCAGCACTGATCCCTCTCA 564
114 sHisAspHisHisAspGluPheCys.LeuMetPro 125
565 CCATGACCATCATCAGATGAGTTTGCCCTTATGCC 599

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.AAC03880

seq_documentation_block:

ID AAC03880 standard; cDNA; 662 BP.

AC AAC03880;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 3878.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

OS Homo sapiens.

XX EP1033401-A2.

PN 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR P-PSDB; AAG03874.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 3878; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

CC mRNAs encoding secreted proteins. An ORF has been identified within the

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

CC derived from 30 different tissues. EST sequences usually correspond

CC mainly to the 3' untranslated region (UTR) of the mRNA because they are

CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

CC well suited for isolating cDNA sequences derived from the 5' ends of

CC mRNAs and even in those cases where longer cDNA sequences have been

CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

CC mRNAs with intact 5' ends and can therefore be used to obtain full length

CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

CC gene therapy and chromosome mapping procedures. They are used to obtain

CC upstream regulatory sequences and to design expression and secretion

XX vectors.

XX Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

alignment_scores:

Quality: 595.50 Length: 128

Ratio: 4.920 Gaps: 1

Percent Similarity: 92.969 Percent identity: 86.719

alignment_block:

US-09-327-750D-32 x AAC03880 ..

Align seg 1/1 to: AAC03880 from: 1 to: 662

1 MetGluSerLysGluLysArgAlaValAsnSerLeuMetGluAsnAl 17

|||||
207 ATGAGTCCAAAGAGGACGCGTTAAACAATCTCATCTCGTGGAAATGT 256
17 asnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
257 CAACACAGGAAATGATGAAAAAGATGAAAAAGAGGACAGTGCCTAATAAG 306
31 LysGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
307 GGGAGCCCTTGCCCTTACCTTTGATGTYTRGTGATCTGTGTGCCTAGA 356
48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
357 GGAATCGTAGCGGTTCGCGCTTAGGAGCCCATCTCTGCAGTAGATG 406
64 PAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluAla 81
407 GGATATGATCATAGCTTGGAGAACACAGGAGGATGAGNAGAAGA 456
81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
457 ATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGAAGCTGAGG 506
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
507 GAAAGCAGTTGATCATAGTCTGCGGCGAGTCAGCACTGACCCCTCTCA 556
114 sHisAspHisHisAspGluPheCysLeuMetPro 125
557 CCATGACCATCATCAGATGAGTTTGCCCTTATGCC 599

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.AAC03879

seq_documentation_block:

ID AAC03879 standard; cDNA; 698 BP.

AC AAC03879;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 3877.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG03873.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

XX diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 3877; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

CC mRNAs encoding secreted proteins. An ORF has been identified within the

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

CC derived from 30 different tissues. EST sequences usually correspond

CC mainly to the 3' untranslated region (UTR) of the mRNA because they are

CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

XX SQ Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:
 Quality: 585.50 Length: 128
 Ratio: 4.920 Gaps: 1
 Percent Similarity: 92.969 Percent Identity: 86.719

alignment_block:

US-09-327-750D-32 x AAC03879 ..

Align seg 1/1 to: AAC03879 from: 1 to: 698

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 243 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCTGGGAAATCT 292
 17 asanGlnGluasn.....GluGluLysGluGlnValAlaAsnLysG 31
 293 CAACGAGGAAATGATGAAAAGATGAAAGGAGCAAGTGCCTAATAAAG 342
 31 lYGlUProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 343 GGGAGCCCTTGGCCCTACCTTGTGATGCTGTGCTGTGCTGTGCTGTG 392
 48 GlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 54
 393 GGAATATCGTAGCGGTTCCCGGTTAGGACGCCCATCTCTGCAGTATAGATG 442
 64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 81
 443 GGATATGATGATAGCTTGGAGCAACACAGGCAAGGATGAGAGAGAGA 492
 81 snMetCluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 493 ATATGGAAGGATTGGGAGGAGGTGAGACGCTGATGGAAGCTGAGG 542
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 114
 543 GAAAGACAGTTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCA 592
 114 sHisAspHisHisAspGluPheCysLeuMetPro 125
 593 CCATGACCATCATGATGATGTTTCNNWATGCC 626

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC10889

seq_documentation_block:

ID AAC10889 standard; cDNA; 692 BP.

XX AC AAC10889;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 14964.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

XX PN EP1033401-A2.

XX

PD 06-SEP-2000.
 XX 21-FEB-2000; 2000EP-0200610.
 XX 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1: SEQ ID 14964; 71pp + CD-ROM; English.
 XX The present sequence is one of a large number of 5' ESTs derived from
 XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
 XX identified within the present sequence. The 5' ESTs were prepared from
 XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 XX sequences usually correspond mainly to the 3' untranslated region (UTR)
 XX of the mRNA because they are often obtained from oligo-dT primed cDNA
 XX libraries. Such ESTs are not well suited for isolating cDNA sequences
 XX derived from the 5' ends of mRNAs and even in those cases where longer
 XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
 XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX They are used to obtain upstream regulatory sequences and to design
 XX expression and secretion vectors.
 XX SQ Sequence 692 BP; 184 A; 151 C; 199 G; 146 T; 12 other;

alignment_scores:
 Quality: 562.50 Length: 125
 Ratio: 4.849 Gaps: 1
 Percent Similarity: 92.800 Percent Identity: 84.800

alignment_block:

US-09-327-750D-32 x AAC10889 ..

Align seg 1/1 to: AAC10889 from: 1 to: 692

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG 20
 246 CAAAGAGACGAGCGTTAAACAATCTCATCTGGGAAATGTCACACAGGA 295
 20 uasn.....GluGluLysGluGlnValAlaAsnLysGlyGluProL 34
 296 AAATGATGAAAAGATGAAAGGAGCAAGTGCCTAATAAAGGGGAGCCCT 345
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 346 TGGCCCTACCTTGTGATGTGTGTAATCTGTGTGCTGTAGAGGAAATCGT 395
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
 396 AGGCGGTCTCCGGTGGAGCCCATCTCTGCAGTATAGATGGGATATGAT 445
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGlu 84
 446 GCATAGGCTTGGAGAACCCACAGGCAAGGATGAGAGAGAGATATGAAA 495
 84 rGleGlyGluGlnValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 496 GGATTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGGGAACAGCAG 545
 101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspH 117
 546 TTGAGTCATAGTCTCGGGGAGCTGACAGCACTGACCCCTCACCATGACCA 595

117 shisaspGluPheCysLeuMetPro 125
 |||||
 596 TCATGATGAGTTTGCNNWATGCC 620

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT: AAC06183

seq_documentation_block:

ID AAC06183 standard; cDNA; 421 BP.

AC AAC06183;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10258.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 10258; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

XX Sequence 421 BP; 110 A; 103 C; 124 G; 82 T; 2 other;

alignment_scores:

Quality: 207.50 Length: 72
 Ratio: 3.402 Gaps: 2
 Percent Similarity: 84.722 Percent Identity: 65.278

alignment_block:

US-09-327-750D-32 x AAC06183

Align seg 1/1 to: AAC06183 from: 1 to: 421

1 MetGluSerLysGluValAsnSerLeuMetGluAsnAl 17

207 ATGGAGTCCAAAGAGGAGGACGAGGCTTAAACAATCTCATCGTGGAAATGT 256

17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLys 31
 :|||||
 257 CAACACGAGAAATGATGAAGAAGATGAAGAAGAGCAAGTGTCTAATAAG 306
 |||||
 31 LysGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||
 307 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGTGCTTAGA 356
 |||||
 48 GlyAsnArgArgArgPheProValArgGlnProIleLeuGln.TyrArgT 64
 |||||
 357 GGAACCCGTAG.CGTTCCGGCTTAGGCACACAGTATTAGAGGCACCGCT 405
 |||||
 64 rPaspIleMethis 68
 :|||
 406 GCCCAGTGCACAT 419

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT: AAC06180

seq_documentation_block:

ID AAC06180 standard; cDNA; 457 BP.

XX AAC06180;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10255.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 10255; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

XX Sequence 457 BP; 120 A; 110 C; 142 G; 84 T; 1 other;

alignment_scores:

Quality: 207.50 Length: 72
 Ratio: 3.402 Gaps: 2

Percent Similarity: 62.037 Percent Identity: 44.444
alignment_block:
US-09-327-750D-32 x AAF21748
Align seg 1/1 to: AAF21748 from: 1 to: 917
19 GlnGluAsnGluGluLysGluValAlaAsnLysGlyGluProLeuAl 35
|||||
306 CAGGAAACGAGAGATGGAGCCCTATCGAAGTGGAGAGAA...GA 352
35 aLeuProLeuAspAlaGlyGlyTyrCysValProArgGlyAsnArgArg 52
|||||
353 CGCCCTTTGGAGAGGTGAAGGCCACCGCTGCAGGAAATCGACGG 402
52 rgPheProValArgGlnProIleLeuGlnTyrArgTyrPaspileMetHis 68
|||||
403 GA...CAGGCTCGCGGACTTGGCCCTAATTTTCATGGGCCATACCCAT 449
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI 85
|||||
450 AGGCAGATCAATGATGGGTGGATGGAGATGATGATGATGATGATGAT 499
85 eclyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeu 102
|||||
500 CATGGAGAGATGAGAGAAATCAGAGAAACTTAGGAGCTCCAGTTGA 549
102 erHisSerLeuArgAlaValSerThrAsp...ProProHisHisAspHis 117
|||||
550 GGAATTGTCTGCTATCTTATGGGGAGCTCTCTAATCACCACCATGACC 599
118 HisAspGluPheCysLeuMetPro 125
|||||
600 CATGATGAATTTTGCCTATGCTT 623

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748
seq_documentation_block:
ID AAF21748 standard: DNA; 917 BP.
AC AAF21748;
DT 27-MAR-2001 (first entry)
XX Human breast and ovarian cancer associated antigen gene SEQ ID 135.
DE
XX Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW nontropic; neurprotection; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS Homo sapiens.
XX
XX WO200055173-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05881.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI: 2000-611515/58.
XX P-PSDB: AAB58845.
XX
XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
XX Claim 1; Page 581-582; 1299pp; English.
XX
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
XX Sequence 917 BP; 228 A; 203 C; 270 G; 211 T; 5 other;

alignment_scores:
Quality: 185.50 Length: 108
Ratio: 2.769 Gaps: 3

gb_hgt:AC068585	+	86.00	105.66	1.e+03	70358	!	AC068585	Homo sapiens chr1
gb_hgt:AC069313	+	86.00	100.00	2.7e+03	129315	!	AC069313	Trypanosoma cruzi
gb_ro:AB041591	+	85.00	134.99	30.49	1657	!	AB041591	Mus musculus brain
gb_hgt:AC013960	+	85.00	102.53	2.1e+03	83821	!	AC013960	Drosophila melan
gb_hgt:AC009346	+	85.00	96.70	4.1e+03	168448	!	AC009346	Drosophila melan

seq_name: qb_pr:AF183416

```

seq_documentation_block:
LOCUS AF183416 642 bp mRNA PRI 02-SEP-2000
DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
mRNA, complete cds.
ACCESSION AF183416
VERSION AF183416.1 GI:9963770
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 642)

```

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1

```

/db_xref="taxon.9606"
/tissue_type="adrenal gland"
3...380
/note="HGR74-h protein"
/codon_start=1
/evidence=not_experimental
/product="ovarian granulosa cell 13.0 kDa protein hGR74 homolog"
/protein_id="AAG09685.1"
/db_xref="GI:9963771"
/translators="MSKCKRAVNSLSMENAOENKEEQVANKGEPLALPLDAGEYC
VPGNRNRRFRVRPQILQYRDMDHRLGEQPARRENMERIGEYVRLMEKLREKQLS
HSLRAVSTDPPHHDDHDFCLMP"
BASE COUNT      204 a 118 c 157 g 163 t
ORIGIN

alignment_scores:
    Quality: 653.00      Length: 125
    Ratio: 5.266
Percent Similarity: 99.200      Percent Identity: 98.400

alignment_block:
US-09-327-750D-32 x AF183416 ..

Align seg 1/1 to: AF183416 from: 1 to: 642

1 MetGluSerIysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
3 ATGGAGTCCAAAGAGAAAGACGACGTAACAGTCFCACGATGGAAATGC 52
|||||
17 aaSngInGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
53 CAACCAAGAAAAATGAAGAAAGGACGAGCTTGCTAATAAAGGGGAGCCCT 102
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
103 TGGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCCCTAGAGGAATCGT 152
|||||
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||

```

```

153 AGCGGTTCCGGTTAGCAGCCCATCTCCAGCTATAGATGGGATATGAT 202
67 tHisArgLeuGluGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
203 GCATAGGCTTGGAGAACACAGGCAAGATGAGAGAGAGATATGGA 252
84 rgileGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
253 GGATTGGGAGGAGGTGAGACAGCTGATGGAAAGCTGAGGGAAGACAG 302
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspH1 117
|||||
303 TTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCACCATGACCA 352
117 sHisAspGluPheCysLeuMetPro 125
|||||
353 TCATGATGAGTTTTCCTTATGCC 377

```

seq_name: gb_pr:AF220189

```

seq_documentation_block:
LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189
VERSION AF220189.1 GI:7689028
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Xiao.H., Song.H., Gao.G., Ren.S., Chen.Z. and Han.Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
REFERENCE
AUTHORS Xiao.H., Song.H., Gao.G., Ren.S., Chen.Z. and Han.Z.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China

```

```

FEATURES
source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67654.1"
/db_xref="GI:7689029"
/translation="MESKEKRAVNSLSMENANQENEEKQVANKGEPLALPLDAGEYC
VRGNRRFRVROPQILOVRMDMHLGEPQARMERNMERIGEVRQLMEKLREKLS
HSLRAVSTDPHPHDDHDFCLMP"
BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN

```

```

alignment_scores:
Quality: 653.00 Length: 125
Ratio: 5.266 Gaps: 0
Percent Similarity: 99.200 Percent Identity: 98.400

```

alignment_block:
US-09-327-750D-32 x AF220189

Align seg 1/1 to: AF220189 from: 1 to: 828

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
167 ATGGAGTCCAAGAGAACGAGCAGTAAACAGTCTCAGCATGGAATATCG 216
17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34

```

```

|||||
217 CAACCAAGAAAATGAAGAAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 266
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
267 TGGCCCTCCCTTTGGATGCTGTGAATACTGTGTGCTAGAGGAATCGT 316
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
317 AGGCGGTTCCGGTTAGCAGCCCATCTCCAGTATAGATGGGATATGAT 366
67 tHisArgLeuGluGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
367 GCATAGGCTTGGAGAACACAGGCAAGATGAGAGAGAGATATGGA 416
84 rgileGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
417 GGATTGGGAGGAGGTGAGACAGCTGATGGAAAGCTGAGGGAAGACAG 466
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspH1 117
|||||
467 TTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCACCATGACCA 516
117 sHisAspGluPheCysLeuMetPro 125
|||||
517 TCATGATGAGTTTTCCTTATGCC 541
seq_name: gb_pr:HS198P4

```

```

seq_documentation_block:
LOCUS HS198P4 35714 bp DNA PRI 16-DEC-2000
DEFINITION Human DNA sequence from clone RP1-198P4 on chromosome Xq22 Contains
a CpG Island, complete sequence.
ACCESSION AL008708
VERSION AL008708.4 GI:4826445
KEYWORDS HPG: CpG Island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 35714)
AUTHORS Corby,N.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clones@sanger.ac.uk
On May 13, 1999 this sequence version replaced GI:4582117.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep/
from the library RPI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://pcypac.med.buffalo.edu/
VECTOR: pcypac2
IMPORTANT: This sequence is not the entire insert of clone

```


seq_name: gb_pr:AF237783

seq_documentation_block: 791 bp mRNA PRI 02-SEP-2000
 LOCUS AF237783 Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete cds.
 DEFINITION Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete cds.
 ACCESSION AF237783
 VERSION AF237783.1 GI:9963898
 KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 791)
 Yang,Q.S., Ying,K., Xie,Y. and Mao,Y.M.
 TITLE A Novel Human X-linked Brain Expressed Protein
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 791)
 Mao,Y.M., Xie,Y., Yang,Q.S., Wu,H., Lin,S. and Ying,K.
 AUTHORS Direct Submission
 TITLE
 JOURNAL Submitted (21-FEB-2000) Genetic Research Institute, Fudan University, 220 Handan Rd., Shanghai 200433, P.R.China

FEATURES
 source
 1..791
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="between DXS990 and DX1059 by RH (Stanford G3)"
 1..791
 /gene="BEX1"
 172..549
 /gene="BEX1"
 /codon_start=1
 /product="brain-expressed protein BEX1"
 /protein_id="AA09752.1"
 /db_xref="GI:9963898"
 /translation="MESKEKLAVNSLSENNQENKEQVANKGEPLPLDAGSYC
 VPGNRRFRVROPILQYRWDMHRLGEPQARMRENNERIGEVRLMEKLEKRLS
 HSLRAVSTDPHDDHDFCLMP"

BASE COUNT 214 a 172 c 218 g 187 t
 ORIGIN

alignment_scores:
 Quality: 639.00 Length: 125
 Ratio: 5.238 Gaps: 0
 Percent Similarity: 97.600 Percent Identity: 96.800

alignment_block:

US-09-327-750D-32 x AF237783 ..

Align seg 1/1 to: AF237783 from: 1 to: 791

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||
 172 ATGGAGTCCAAAGAGAAACTAGCAGTAAACAGTCTCAGCATGGAATGC 221
 17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
 |||||
 222 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 271
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 |||||
 272 TGGCCCTCCCTTTGGATCTGGTGAATCTGTGTGCTAGAGGAATCGT 321
 51 ArgArgPheProValArgLysGlnProLeuGlnTyrArgTyrAspIleMe 67
 |||||
 322 AGCGGTTCCGCGTTAGCAGCCCATCTCCAGCTATAGATGCGATATGAT 371
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGlu 84
 |||||
 372 GCATAGGCTTGAGAACACACAGGCAAGGATGAGAGAGAGAATATGNA 421

84 rgIleGlyGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 |||||
 422 GGATTGGGAGGGGTGAGACAGCTGATGGAAGAGCTGAGGAAGACGAG 471
 101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspH 117
 |||||
 472 TTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCACCATGACCA 521
 117 sHisAspGluPheCysLeuMetPro 125
 |||||
 522 TCATGATGAGTTTGCCTTATGCC 546

seq_name: gb_pr:AF251053

seq_documentation_block:
 LOCUS AF251053 744 bp mRNA PRI 15-APR-2001
 DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
 ACCESSION AF251053
 VERSION AF251053.1 GI:13625167
 KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 744)
 AUTHORS Mao,Y., Xie,Y., Zhou,Z., Zhao,W., Zhao,S., Wang,W., Huang,Y.,
 Wang,S., Tang,R., Chen,X. and Wu,C.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R. China

FEATURES
 source
 1..744
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 101..487
 /codon_start=1
 /product="X-linked protein"
 /protein_id="AAK34943.1"
 /db_xref="GI:13625168"
 /translation="MESKEERALNNLIVENQENDEKEQVANKGEPLPLNYS
 EYCVPRGNRRFRVROPILQYRWDMHRLGEPQARMRENNERIGEVRLMEKLEK
 QLSHSLRAVSTDPHDDHDFCLMP"

BASE COUNT 222 a 134 c 208 g 180 t
 ORIGIN

alignment_scores:
 Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x AF251053 ..

Align seg 1/1 to: AF251053 from: 1 to: 744

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||
 101 ATGGAGTCCAAAGAGAGAGAGCGTTAAACAATCTCATCGTGGAAATGT 150
 17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 |||||
 151 CAACCAAGGAATGATGAAAAGAGATGAAAAGAGAGCAAGTTGCTAATAAG 200
 31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||
 201 GGGAGCCCTTGGCCCTACCTTTGATGTTAGTGAATACTGTGCTCCTAGA 250
 48 GlyAsnArgArgPheProValArgGlnProLeuGlnTyrArgTyr 64
 |||||
 251 GGAACCGCTAGCGGTTCCGGCTTAGCAGCCCATCTCTGCGAGTATAGATG 300

64 pspilleMethHsArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
 301 GGACATAATGCTAGGCTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 350
 81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 351 ATATGAAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGG 400
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 114
 401 GAAAGACAGTTGAGTCTAGTCTCGGCGCAGTCACTGATCCCTCA 450
 114 shIsAspHisAspGluPheCysLeuMetPro 125
 451 CCATCACCACATCAGATGATTTTGCCTTATGCC 484

seq_name: gb_pat:AX078272

seq_documentation_block:

LOCUS AX078272 898 bp DNA PAT 22-FEB-2001
 DEFINITION Sequence 76 from Patent WO0107471.

ACCESSION AX078272

VERSION AX078272.1 GI:13157963

KEYWORDS

SOURCE

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 898)

AUTHORS Hillman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,

Animai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.

Cell cycle and proliferation proteins

JOURNAL Patent: WO 0107471-A 76 01-FEB-2001;

Incyte Genomics, Inc. (US)

FEATURES

Location/Qualifiers

1..898

/organism="Homo sapiens"

/db_xref="taxon:9606"

/notes="Incyte ID No: 5664154CB1"

BASE COUNT 250 a 186 c 251 g 211 t

ORIGIN

alignment_scores:

Quality: 602.50 Length: 128

Ratio: 4.898 Gaps: 1

Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x AX078272 ..

Align seg 1/1 to: AX078272 from: 1 to: 898

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnA 17
 254 ATGAGTCCAAAGAGAGAGGAGCGCTTAACAACTCATCTCGTGGAAATGT 303
 17 asnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 304 CAACAGGAAATATGATGAAAAGATGAAAGAGGACAAAGTTGCTAATAAG 353
 31 lyluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 354 GGAGGCCCTGGCCCTACTTGTGAATGTAGTACTGTGTCCTAGG 403
 48 GlyAsnArgArgPheProValArgGlnProLeuGlnTyrArgTr 64
 404 GGAACCGTAGGCGGTTCGCGGTAGGAGCCCATCTCGTCAATAGATG 453
 64 pspilleMethHsArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
 454 GGACATAATGCTAGGCTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 503

81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 504 ATATGAAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGG 553
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 114
 554 GAAAGACAGTTGAGTCTAGTCTCGGCGCAGTCACTGATCCCTCA 603
 114 shIsAspHisAspGluPheCysLeuMetPro 125
 604 CCATCACCACATCAGATGATTTTGCCTTATGCC 637

seq_name: gb_pr:HSV870H8

seq_documentation_block:

LOCUS HSV870H8 31321 bp DNA PRI 23-NOV-1999

DEFINITION Human DNA sequence from cosmid V870H8, between markers DXS366 and DXS87 on chromosome X contains ESTs.

ACCESSION 270233

VERSION 270233.1 GI:1235542

KEYWORDS

SOURCE

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 31321)

AUTHORS Whiteley,M.

TITLE Direct Submission

JOURNAL Submitted (19-MAR-1995) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1RQ, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT IMPORTANT: This sequence is not the entire insert of clone V870H8.

It may be shorter because we only sequence overlapping sections

once, or longer because we arrange for a small overlap between

neighbouring submissions.

The true left end of clone V870H8 is at 1 in this sequence. The

true left end of clone V693A8 is at 31221.

V870H8 is from the human chromosome X-specific cosmid library.

FEATURES

Location/Qualifiers

1..31321

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="X"

/clone="GHC-870H8"

/clone_lib="SCCV"

475..701

/note="L1 element fragment"

802..988

/note="L1 element fragment"

2533..4138

/note="match: multiple ESTs"

4189..4246

/note="29 copies of 2 mer 91 % conserved"

7496..7874

/note="L1 element fragment"

8022..8108

/note="L1 element fragment"

8302..8493

/note="L1 element fragment"

8585..8671

/note="L1 element fragment"

8769..9010

/note="L1 element fragment"

9259..9384

/note="L1 element fragment"

9624..9731

/note="MLTIC element fragment"

9781..9948

/note="MLTIC element fragment"

9890..9948

/note="MLT1B element fragment"

10418..10592

```

/partial
/note="Alu repeat: matches 298..108 of consensus"
10593..10688
/partial
/note="Alu repeat: matches 96..1 of consensus"
10698..10781
/note="L1 element fragment"
11112..11403
/note="Alu repeat: matches 1..308 of consensus"
11922..12004
/note="L1 element fragment"
12062..12308
/note="L1 element fragment"
12423..12707
/partial
/note="Alu repeat: matches 302..1 of consensus"
12710..12916
/note="L1 element fragment"
12979..13081
/note="L1 element fragment"
13473..13532
/note="MT2A1 element fragment"
13871..13928
/note="L1 element fragment"
18180..18417
/note="L1 element fragment"
18778..19056
/note="Alu repeat: matches 1..308 of consensus"
22279..22350
/note="L1 element fragment"
22449..22995
/note="L1 element fragment"
23239..23374
/note="L1 element fragment"
23371..23606
/note="MT1B element fragment"
23395..23541
/note="MT1A element fragment"
23705..23778
/note="MT1A element fragment"
23805..23960
/note="L1 element fragment"
24011..24541
/note="L1 element fragment"
24638..24770
/partial
/note="Alu repeat: matches 150..1 of consensus"
24829..30975
/note="L1 element fragment"
BASE COUNT 9462 a 6147 c 6129 g 9583 t
ORIGIN

alignment_scores:
  Quality: 602.50      Length: 128
  Ratio: 4.898        Gaps: 1
  Percent Similarity: 96.094  Percent Identity: 89.062

alignment_block:
US-09-327-750d-32 x HSV870H8
Align seg 1/1 to: HSV870H8 from: 1 to: 31321

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
3508 ATGGAGTCAAGAGGACGAGGCTTAACAATCTCTGCTGGAAAATG 3557
|||||
17 asnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
|||||
3558 CAACGAGGAATGATGAAAAGATGAAAAGGACGAGTGTCTATTAAG 3607
|||||
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||

```

```

3608 GGGAGCCCTTGCCCTACCTTTGAATGTAGTGAATAGTGTGCTAGTA 3657
48 GlyAsnAlaGArgProValArgGlnProIleLeuGlnTyrArgTr 64
|||||
3658 GGAACCCGTAGCGGTTCCGGTTAGGACGCCATCTCGCAGTAGATG 3707
64 pasPileMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
|||||
3708 GGACATAATGTCATAGCTTGAGAGCCACAGGCAAGGATGAGAGAGAGA 3757
81 snMetGluArgIleGlyGluGlnValArgGlnLeuMetGluLysLeuArg 97
|||||
3758 ATATGGAAGGATTGGGAGGAGCTGACACAGCTGATGGAAGAGCTGAGG 3807
98 GluLysGluLeuSerHisSerLeuArgAlaValSerThrAspProProH 114
|||||
3808 GAAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCAGCTGATCCCTCA 3857
114 shisAspHisHisAspGluPheCysLeuMetPro 125
|||||
3858 CCATGACCATCAGATGAGTTTGCCTTATGCC 3891
seq_name: gb_pr:AL133348
seq_documentation_block:
LOCUS AL133348 40584 bp DNA PRI 06-MAR-2000
DEFINITION Human DNA sequence from clone RPL-79p11 on chromosome Xq21.32-22.1.
Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed X-linked protein 2), ESTs, STSS, GSSs and a
putative CpG island, complete sequence.
ACCESSION AL133348
VERSION AL133348.8 GI:7076386
KEYWORDS HTG; Bex2; CpG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40584)
AUTHORS Wilson, S.
TITLE Direct Submission
JOURNAL
COMMENT
On Feb 24, 2000 this sequence version replaced gi:6997869.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RPL-79p11 is from the library RPL-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RPL-79p11 it may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

```

The true left end of clone LLOXNC01-105G4 is at 40485 in this sequence. The true right end of clone LLOXNC01-177E8 is at 100 in this sequence.

FEATURES

```

source
    Location/Qualifiers
    1..40584
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="X"
    /map="q21.32-22.1"
    /clone="RPL-79P11"
    /clone_lib="RPC1-1"
    19..292
    /note="AluJo repeat: matches 3..275 of consensus"
    repeat_region
    470..1284
    /note="LIMB5 repeat: matches 5318..6157 of consensus"
    repeat_region
    1295..1898
    /note="LIPAL5 repeat: matches 5556..6157 of consensus"
    repeat_region
    1899..2345
    /note="LIPAL5 repeat: matches 4922..5370 of consensus"
    repeat_region
    2323..2698
    /note="LIMB5 repeat: matches 4960..5343 of consensus"
    repeat_region
    2927..3065
    /note="MER7A repeat: matches 164..305 of consensus"
    repeat_region
    3655..9795
    /note="LIPAL4 repeat: matches 2..6142 of consensus"
    repeat_region
    9866..9982
    /note="FLAM_C repeat: matches 6..124 of consensus"
    repeat_region
    9989..10254
    /note="L1 repeat: matches 4027..4293 of consensus"
    repeat_region
    10253..10852
    /note="LIM4 repeat: matches 4755..5333 of consensus"
    repeat_region
    10853..11260
    /note="MT1B repeat: matches 3..390 of consensus"
    repeat_region
    11261..11474
    /note="LIM4 repeat: matches 4538..4755 of consensus"
    repeat_region
    11480..11563
    /note="MSTB repeat: matches 1..80 of consensus"
    repeat_region
    11564..12475
    /note="L1 repeat: matches 3039..3971 of consensus"
    misc_feature
    14463..14861
    /note="match: GSS: Em:AQ729215"
    repeat_region
    15578..15854
    /note="AluJo repeat: matches 6..289 of consensus"
    misc_feature
    complement(16007..16556)
    /note="match: GSS: Em:AQ772810"
    repeat_region
    16197..16569
    /note="LIM4 repeat: matches 4267..4627 of consensus"
    repeat_region
    16629..16714
    /note="LIM4c repeat: matches 1578..1666 of consensus"
    repeat_region
    16902..17133
    /note="LIM4 repeat: matches 0..242 of consensus"
    repeat_region
    17536..17728
    /note="MIR repeat: matches 1..191 of consensus"
    repeat_region
    17746..18323
    /note="L2 repeat: matches 1599..2175 of consensus"
    repeat_region
    18706..18755
    /note="LIMB3 repeat: matches 6034..6083 of consensus"
    repeat_region
    18756..19236
    /note="MER6B repeat: matches 1..486 of consensus"
    repeat_region
    19237..21108
    /note="MER66-internal repeat: matches 4808..6676 of consensus"
    repeat_region
    21109..21163
    /note="MT2B repeat: matches 391..445 of consensus"
    repeat_region
    21164..21555
    /note="MER6B repeat: matches 3..391 of consensus"
    repeat_region
    21557..21662
    /note="LIMB3 repeat: matches 6080..6185 of consensus"
    repeat_region
    21665..21761
    /note="LIM repeat: matches 5299..5392 of consensus"
    repeat_region
    21762..21929
    /note="LIMB8 repeat: matches 5902..6063 of consensus"
    repeat_region
    21930..22212
    /note="Alusg repeat: matches 3..284 of consensus"
    22213..22321
    /note="LIMB8 repeat: matches 6063..6173 of consensus"
    22333..22740
    /note="LIMD2 repeat: matches 5867..6288 of consensus"
    22744..23094
    /note="LIMD repeat: matches 2..342 of consensus"
    23218..23524
    /note="Alusg repeat: matches 1..305 of consensus"
    23603..23947
    /note="LIM4 repeat: matches 4916..5184 of consensus"
    23948..24218
    /note="AluJo repeat: matches 1..282 of consensus"
    24219..24510
    /note="LIM4 repeat: matches 4601..4916 of consensus"
    24508..24621
    /note="LIM4 repeat: matches 4374..4488 of consensus"
    24650..25077
    /note="MT1C repeat: matches 1..466 of consensus"
    25078..25914
    /note="LIM4 repeat: matches 3497..4346 of consensus"
    25920..26741
    /note="LIM4 repeat: matches 2202..3065 of consensus"
    26762..26906
    /note="LIPAL7 repeat: matches 6001..6145 of consensus"
    26907..27135
    /note="LIPAL5 repeat: matches 5674..5902 of consensus"
    27266..27509
    /note="LIM4 repeat: matches 2164..2080 of consensus"
    27557..27769
    /note="LIMC repeat: matches 253..460 of consensus"
    28120..28167
    /note="L2 repeat: matches 2703..2750 of consensus"
    29043..29349
    /note="MER5B repeat: matches 1..341 of consensus"
    30389..30452
    /note="16 copies 4 mer tata 89% conserved"
    polyA_site
    complement(30502)
    /gene="dj79p11.1"
    mRNA
    complement(join(30502..31137,31514..31589,32016..>32134))
    /gene="dj79p11.1"
    /note="match: ESTs: Em:AA317587 Em:AA358632 Em:AI214048 Em:AI292470 Em:AA319216 Em:AA317625 Em:D81392 Em:AA317418 Em:AA366528 Em:CI5186 Em:AI291270 Em:AA160660 Em:AA612605 Em:AW160832 Em:AA384405 Em:AA160657 Em:AA081744 Em:AI291126 Em:CI5591 Em:AA384268 Em:AI929703 Em:AI302026 Em:W19547 Em:AI141727 Em:AA946933 Em:AI141583 Em:H61106 Em:W60581 Em:AI826617 Em:AA706545 Em:AA64771 Em:AA862327 Em:W68380 Em:AA641652 Em:AA384269 Em:AA527696 Em:W48778 Em:AA633989 Em:AI659486 Em:AW167301 Em:AI929428 Em:AW379411 Em:CI5948 Em:AI745669 Em:AW024587 Em:AW379207; match: proteins: Tr:Q9Y516 Tr:Q9Y517"
    /evidence=not_experimental
    /product="dj79p11.1 (novel brain-expressed x-linked protein 2)"
    complement(30502..32134)
    /gene="dj79p11.1"
    30507..30988
    /note="match: STS: Em:G24641; match: STS: Em:T86927"
    30508..30852
    /note="match: STS: Em:R43117; match: STS: Em:G23964"
    complement(30518..30523)
    /gene="dj79p11.1"
    30541..30922
    /note="match: STS: Em:H61107"
    31343..32262
    /note="CpG island"
    /evidence=not_experimental
    31647..31730
    /note="42 copies 2 mer cc 64% conserved"
    complement(32655..33316)
    /note="match: GSS: Em:AQ489303"
    33027..34514

```

/notes="LIME2 repeat: matches 4631. .6155 of consensus"
 34528. .34786
 /notes="LIME3A repeat: matches 4404. .4664 of consensus"
 34787. .35085
 /notes="AluSg repeat: matches 1. .300 of consensus"
 35086. .35412

alignment_scores:
 Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x AL133348/rev ..

Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||.....
 31132 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCGTGGAAAATGT 31083

17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 |||||.....

31082 CAACCCAGGAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTTAATAAG 31033

31 lYgluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||.....
 31032 GGGAGCCCTGGCCCTACCTTGAATGTTAGTGAATCTGTGTCCTAGA 30983

48 GlyAsnArgArgPheProValArgGlnProLeuGlnTyrArgTr 64
 |||||.....

30982 GGAACCCGTAGCGGTTCCGCGTTAGGACGCCCATCTCGAGTATAGTG 30933

64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 81
 |||||.....

30932 GGACATATGCTAGGCTTGAGAGCCACAGGCAAGGATGAGAGAGAGA 30883

81 sMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
 |||||.....

30882 ATATGGAAGGATTGGGAGAGGTGAGACAGCTGATGGAAAAGCTGAGG 30833

98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 114
 |||||.....

30832 GAAAGCAGTTGAGTCATAGTTTCGGGCGAGTCAGCACTGATCCCCCTCA 30783

114 sHisAspHisHisAspGluPheCysLeuMetPro 125
 |||||.....

30782 CCATGACCATCAGTATGAGTTTGCCTTATGCC 30749

seq_name: gb_htg:HSU80B1

seq_documentation_block:

LOCUS HSU80B1 41029 bp DNA HTG 10-JUL-2001
 DEFINITION Homo sapiens chromosome X clone LLOXNC01-80B1 map q22.1-22.3, ***
 SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AL022169

VERSION AL022169.3 GI:13276704

KEYWORDS HTG; HTGS_PHASE2; HTGS_CANCELLED.

SOURCE human

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 41029)

Bird.C.

Direct Submission

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Mar 12, 2001 this sequence version replaced gi:12750927.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: CU80B1
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator ABI; 19% of reads
 Consensus quality: 41029 bases at least Q40
 Consensus quality: 41023 bases at least Q30
 Consensus quality: 41027 bases at least Q20
 Insert size: 41029; sum-of-contigs
 Insert size: 51751; 0.6% error; agarose-fp
 Quality coverage: 9.67x in Q20 bases; sum-of-contigs Quality
 coverage: 7.67x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

Source

1. 41029
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="q22.1-22.3"
 /clone="LLOXNC01-80B1"
 /clone_lib="LLOXNC01"
 misc_feature 1. 41029
 /note="assembly_fragment:00890"
 BASE COUNT 13249 a 7796 c 7781 g 12203 t
 ORIGIN

alignment_scores:

Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x HSU80B1 ..

Align seg 1/1 to: HSU80B1 from: 1 to: 41029

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

18412 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCGTGGAAAATGT 18461

17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 |||||.....

18462 CAACCCAGGAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTTAATAAG 18511

31 lYgluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||.....

18512 GGGAGCCCTTGCCCTACCTTGAATGTTAGTGAATCTGTGTCCTAGA 18561

48 GlyAsnArgArgPheProValArgGlnProLeuGlnTyrArgTr 64
 |||||.....

18562 GGAACCCGTAGCGGTTCCGCGTTAGGACGCCCATCTCGAGTATAGTG 18611

64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 81
 |||||.....

18612 GGACATATGCTAGGCTTGAGAGCCACAGGCAAGGATGAGAGAGAGA 18661

81 sMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
 |||||.....

18662 ATATGGAAGGATTGGGAGGAGGTGAGACACCTGATGGAAAAGCTGAGG 18711

98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 114
 |||||.....

18712 GAAAAGCAGTTGAGTCAATAGTCTCGGGCAGTCAGCACTGATCCCCCTCA 18761

114 sHisAspHisHisAspGluPheCysLeuMetPro 125

```

18762 CCATGACCATCAGGATGAGTTTCCTTATGCCC 18795
seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS AF097439 785 bp mRNA 13-APR-1999
DEFINITION Mus musculus brain expressed X-linked protein 2 (Bex2) mRNA,
complete cds.
ACCESSION AF097439
VERSION AF097439.1 GI:4580591
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 785)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum:[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE 2 (bases 1 to 785)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
Source 1..785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:W48832"
/chromosome="X"
/map="near Plp"
/dev_stage="embryo; 15.5 dpc"
1..785
/gene="Bex2"
139..528
/gene="Bex2"
/codon_start=1
/product="brain expressed X-linked protein 2"
/protein_id="AAD24430.1"
/db_xref="GI:4580592"
/translation="MESKVEQGVKNLNMENHDSKEEKEKPODASKRDPVALPREA
GDYVPRGRRFRVROPVIAHWRDLQVYGEPOGRMEENVQRFQDGVRLMEKLE
QLSHSLRAVSTDPHHDHDFCLMP"
BASE COUNT 212 a 163 c 235 g 175 t
ORIGIN

alignment_scores:
Quality: 496.00 Length: 129
Ratio: 4.168 Gaps: 2
Percent Similarity: 92.248 Percent Identity: 72.093

alignment_block:
US-09-327-750D-32 x AF097439 ..
Align seg 1/1 to: AF097439 from: 1 to: 785
1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
139 ATGGAGTCCAAAAGTGCAAGAGCGCTGAAATCTCAACATGGAGATGA 188
17 aasngGluAsnGluLysGluLysGluLysGluLysGluLysGluLysG 31
189 CCATCAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 238
31 lyGluProLeu...AlaLeuProLeuAspAlaGlyGluTyrCysValPro 46
:::|||||: |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

239 GGGATCCGATTGTCGCCCTTCGAGAGCTGAGAGCTACTAGTGCCT 288
47 ArgGlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrAl 63
289 AGAGGAGGTCCGAGCGGTTCCGGGTCGGCAGCCCATCGTGGCTACAG 338
63 gtrpAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGlu 80
339 ATGGGACCTGATGCATAGGTTGGGAGCCCGAGGAGGATGAGAGAGG 388
80 luasnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 96
389 AGAACGTACAGAGGTTGGGATGATGTGACACAGCTCATGGAGAAGCTG 438
97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 113
439 AGGGAAGCAGCTGAGCAGCAGCTCGGGCGGTAGCAGCTGACCGGCC 488
113 OHHisAspHisHisAspGluPheCysLeuMetPro 125
489 TCATCATGACCAACCATGATGAGTTTGCCTTATGCCC 525
seq_name: gb_ro:AF097437

seq_documentation_block:
LOCUS AF097437 2269 bp DNA 05-AUG-1999
DEFINITION Mus musculus Bex1 protein (Bex1) gene, complete cds.
ACCESSION AF097437
VERSION AF097437.1 GI:5702155
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 2269)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum:[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE 2 (bases 1 to 2269)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4030, Australia
FEATURES
Source 1..2269
/organism="Mus musculus"
/strain="129SVJ"
/db_xref="taxon:10090"
/chromosome="X"
/map="near Plp"
/tissue_type="liver"
/join(409..509,866..942,1263..1873)
/gene="Bex1"
/product="Bex1 protein"
409..1873
/gene="Bex1"
/note="expressed in brain; X-linked"
1268..1654
/gene="Bex1"
/codon_start=1
/product="Bex1 protein"
/protein_id="AAD47168.1"
/db_xref="GI:5702156"
/translation="MESKDOGVKNLNMENHDSKEEKEKPODASKRDPVALPREA
KNCAPRGRRFRVROPVIAHWRDLQVYGEPOGRMEENVQRFQDGVRLMEKLE
QLSHSLRAVSTDPHHDHDFCLMP"
BASE COUNT 563 a 554 c 705 g 447 t
ORIGIN

```

```

alignment_scores:
  Quality: 448.50      Length: 129
  Ratio: 3.934         Gaps: 3
  Percent Similarity: 88.372  Percent Identity: 68.217

alignment_block:
  US-09-327-750D-32 x AF097437
  Align seg 1/1 to: AF097437 from: 1 to: 2269

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1268 ATGGAGTCCAAAGAT...CAAGCGTGAAAAATCTCAACATGAGAGATGA 1314
17 AsnGlnGluAsnGluLysGlu.....GlnValAlaAsnLysG 31
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1315 CCATCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1364
31 LyGluPro...LeuAlaLeuProLeuAspAlaGlyGluTyrCysValPro 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1365 GGGAGCCAGCTGTGCGCCCTGACCTCGAGGCTGGCAAAACTGTGCACCT 1414
47 ArgGlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrAr 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1415 AGAGGAGGTGCGAGCGGTTCCGGGTTCGGCAGCCCATCGCTCACTATAG 1464
63 gTrpAspIleMethHisArgLeuGluGluProGlnAlaArgMetArgGluG 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1465 ATGGGACCTGATGACAGAGGTTGGGAGGCCCCAGGAGGAGGAGGAGGAG 1514
80 LuAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1515 AGAAGCTACAGAGGTTTGGGGGTGATGTGAGCAGCTCATGAGAAGCTG 1564
97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1565 AGGNAAGGACGTGACCCACACCTTGGGGGGTTAGCACTGACCCGCC 1614
113 oHisHisAspHisHisAspLysPheCysLeuMetPro 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1615 TCATCATGACCACCATGATGAGTTTGCCTCATGCC 1651

seq_name: gb_ro:AF097438

seq_documentation_block:
  LOCUS AF097438 835 bp mRNA ROD 11-APR-1999
  DEFINITION Mus musculus brain expressed x-linked protein 1 (Bex1) mRNA,
  complete cds.
  ACCESSION AF097438
  VERSION AF097438.1 GI:4580589
  KEYWORDS
  SOURCE house mouse.
  ORGANISM Mus musculus.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  Bex1, a gene with increased expression in parthenogenetic embryos,
  is a member of a novel gene family on the mouse X chromosome
  JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
  MEDLINE 99172070
  REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
  May;8(5):943]]
  REFERENCE 2 (bases 1 to 835)
  AUTHORS Brown, A.L. and Kay, G.F.
  TITLE Direct Submission
  JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
  Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
  FEATURES
    source
      1..835
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /chromosome="X"

```

```

/map="near Plp"
/tissue_type="blastocyst"
1..835
/gene="Bex1"
205..591
/gene="Bex1"
/codon_start=1
/product="brain expressed X-linked protein 1"
/protein_id="AAD24429.1"
/db_xref="GI:4580590"
/translucion="MESKDOGKVLNMMENDHKKKEEKPODTRREPAVALISEAG
KNCAPGGRRFRVRQPTAHVENDLMORVGPQGRMRREENVORFGDVRQIMKELRER
QLSHSLRAVSTDPHPHDDHDEFLMP"
BASE COUNT 238 a 184 c 246 g 167 t
ORIGIN

alignment_scores:
  Quality: 446.50      Length: 129
  Ratio: 3.917         Gaps: 3
  Percent Similarity: 88.372  Percent Identity: 68.217

alignment_block:
  US-09-327-750D-32 x AF097438
  Align seg 1/1 to: AF097438 from: 1 to: 835

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 ATGGAGTCCAAAGAT...CAAGCGTGAAAAATCTCAACATGAGAGATGA 251
17 AsnGlnGluAsnGluGluLysGlu.....GlnValAlaAsnLysG 31
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
252 CCATCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
31 LyGluPro...LeuAlaLeuProLeuAspAlaGlyGluTyrCysValPro 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
302 GGGAGCCAGCTGTGCGCCCTGATCTCCGAGGCTGGCAAAACTGTGCGCCT 351
47 ArgGlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrAr 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
352 AGAGGAGGTGCGAGCGGTTCCGGGTTCGGCAGCCCATCGCTCACTATAG 401
63 gTrpAspIleMethHisArgLeuGluGluProGlnAlaArgMetArgGluG 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
402 ATGGGACCTGATGACAGAGGTTGGGAGGCCCCAGGAGGAGGAGGAGGAGG 451
80 LuAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
452 AGAAGCTACAGAGGTTTGGGGGTGATGTGAGCAGCTCATGAGAAGCTG 501
97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
502 AGGNAAGGACGTGACCCACACCTTGGGGGGTTAGCACTGACCCGCC 551
113 oHisHisAspHisHisAspLysPheCysLeuMetPro 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
552 TCATCATGACCACCATGATGAGTTTGCCTCATGCC 588

seq_name: gb_ro:AF051347

seq_documentation_block:
  LOCUS AF051347 838 bp mRNA ROD 01-OCT-1998
  DEFINITION Mus musculus REX-3 mRNA, complete cds.
  ACCESSION AF051347
  VERSION AF051347.1 GI:3510642
  KEYWORDS
  SOURCE house mouse.
  ORGANISM Mus musculus.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  REFERENCE 1 (bases 1 to 838)
  AUTHORS Faria, T.N., LaRosa, G.J., Willen, E., Liao, J. and Gudas, L.J.

```

TITLE Characterization of genes which exhibit reduced expression during the retinoic acid-induced differentiation of F9 teratocarcinoma cells: Involvement of cyclin D3 in RA-mediated growth arrest Mol. Cell. Endocrinol. 143 (1-2), 155-166 (1998) 99021197

JOURNAL MEDLINE

REFERENCE 2 (bases 1 to 838)

AUTHORS Faria, T.N., LaRosa, G., Wilen, E., Liao, L. and Gudas, L.J.

TITLE Direct Submission

JOURNAL Pharmacology, Cornell University Medical College, 1300 York Avenue, New York, NY 10021, USA

FEATURES

source 1..838

1..838

229..579

/note="retinoic acid reduced gene expression"

/codon_start=1

/product="REX-3"

/protein_id="AAC61929.1"

/db_xref="GI:3510643"

/translation="MENDHOKKEEKEKPOQIRREPAVALISEAGKNCAPRGGRRRF RVQPIAHVMDLMQVGEPOGRMEENVQFGDVRQLMEKLRQLSHSLRAVSTD PPHDHDHDFCLMP"

BASE COUNT 238 a 181 c 251 g 168 t

ORIGIN

alignment_scores:

Quality: 446.50 Length: 129

Ratio: 3.917 Gaps: 3

Percent Similarity: 88.372 Percent Identity: 68.217

alignment_block:

US-09-327-750D-32 x AF051347

Align seg 1/1 to: AF051347 from: 1 to: 838

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

193 ATGGAGTCCAAAGAT...CAGGCGTGAAATCTCAACATGGAGAATGA 239

17 aAsnGlnGluAsnGluLysGlu.....GlnValAlaAsnLysG 31

240 CCATCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 289

31 lylGluPro...LeuAlaLeuProLeuAspAlaGlyGluTyrCysValPro 46

290 GGGAGGCGGAGTGTGGCCCTGTGCTCCGAGGCTGCGCAAAACTGTGCGCCT 339

47 ArgGlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrAr 63

340 AGAGGAGGTGCGAGGCGGTTCGGGTTCGGAGCCCATCGCTACTATAG 389

63 gtrPaspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgLug 80

390 ATGGGACCTGATGACGAGGTTGGGGAGCCCGGAGGAGGAGGAGGAGG 439

80 luAsnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 96

440 AGAACGTACAGAGGTTGGGGGTGATGTGACAGACGCTCATGAGAGCTG 489

97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 113

490 AGGAAAGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 539

113 oHisHisAspHisAspGluPheCysLeuMetPro 125

540 TCATCATGACACCATCATGATGATGATGATGATGATGATGATGATG 576

seq_name: gb_ro:BC003254

seq_documentation_block:

LOCUS BC003254 2901 bp mRNA ROD 20-FEB-2001

DEFINITION Mus musculus. Similar to dentatorubral pallidolysian atrophy, clone MGC:5758, mRNA, complete cds.

ACCESSION BC003254

VERSION BC003254.1 GI:13096906

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 2901)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <http://www.hgsc.bcm.tmc.edu/cdna/> Contact: villalobebcm.tmc.edu. Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 9 Row: 0 Column: 6.

FEATURES

Location/Qualifiers

1..2901

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="MGC:5758 IMAGE:3500522"

/tissue_type="Mammary tumor: C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."

/clone_lib="NCI-CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

755..2359

/codon_start=1

/product="Similar to dentatorubral pallidolysian atrophy"

/protein_id="AAH03254.1"

/db_xref="GI:13096907"

/translation="MVFQELLPLGLPKPPHVTPLHPRAPKEHPTTSHOSLOSP PPQFLPSPPAATGPTTATQIQEPAEYEPPEPPVPAKSPPPKVVDPVS HASQARFNKHLDRGFNSCARSDLYFVLEGSKLAKRADLVKVRREARQAREEKE REREREKERELERSVKLAQEGRAVPCPSLGPVPHRPFPFPGSAVATVPP YLGPDTPALRTLSEYARPHVSPGNRNHPYVPLGAVDPGLGYNVPALYSSDPAARE RERAREEDLDRLKPGFEVKPSELEPLHGVPGLDPPFRHGGALOPGPPGLHPFP FPHSHOHSHHSHLHQDDAIHAASVSHPLDPLASGSHLRIPIYPAGTLEPNLLPH PLHENEVLRLHQLFAAPYRDLPLASAPMSAAHQLOQAHAQSAELORLALQEQOOLHHAH HPLHSVPLPAQEDYYSHLKESDKPL"

BASE COUNT 623 a 958 c 791 g 529 t

ORIGIN

alignment_scores:

Quality: 384.50 Length: 119

Ratio: 3.697 Gaps: 3

Percent Similarity: 87.395 Percent Identity: 65.546

alignment_block:

US-09-327-750D-32 x BC003254

Protocol:
Template: 10 ng

OM of: US-09-327-750D-33 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US05327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINWATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.500 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-DELEXPT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USPR=US09327750 -CGN1.1.5654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-33

Query length: 128

Database: EST*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est1:AL022932	+	646.00	1047.10	3.7e-49	551	AL022932 v8130b28 Beddington mouse
gb_est1:AW536404	-	646.00	1045.88	4.3e-49	707	AW536404 G0104A11-3 NIA Mouse
gb_est2:BG064920	-	643.00	1040.25	8.9e-49	642	BG064920 H3025D01-3 NIA Mouse
gb_est2:BG277659	+	633.00	1025.69	5.7e-48	575	BG277659 uk47c11.y1 Soares_NMMA
gb_est2:BG870503	+	633.00	1023.51	7.6e-48	756	BG870503 G02791422F1 NCI_CGAP_S
gb_est1:BE654459	+	615.00	997.65	2.1e-46	499	BE654459 UI-M-AJ1-aha-f-10-0-UT
gb_est1:BE536974	+	599.00	970.56	6.8e-45	577	BE536974 G0110H10-3 NIA Mouse
gb_est2:BF607762	+	554.00	895.17	1.1e-40	785	BF607762 MYL_000704 Mouse 9-day
gb_est2:BG280077	+	553.00	895.87	1.1e-40	587	BG280077 uk47c11.y1 Soares_NMMA
gb_est2:BF068209	+	516.00	835.09	2.4e-37	650	BF068209 MYL_001088 Mouse 9-day
gb_est1:BE291071	+	505.50	818.94	1.9e-36	583	BE291071 G0108G31F1 NCI_CGAP_N
gb_gss:AE936393	-	493.50	798.78	2.5e-35	637	AE936393 2M0193L05F1 Mouse 10kb
gb_est2:BG669326	+	489.50	795.78	3.7e-35	412	BG669326 DRNAEB11 Rat DRG Libr
gb_est2:BI183527	-	480.50	776.95	4.1e-34	701	BI183527 UNL-P-FN-by-f-07-0-UNI
gb_est1:AW536634	+	472.00	765.12	1.9e-33	549	AW536634 G0106H08-3 NIA Mouse
gb_est2:BG655001	+	471.50	764.34	2.1e-33	547	BG655001 1b44f06.y1 HR85 islet
gb_est2:BG472167	+	471.50	763.17	2.4e-33	634	BG472167 G02513894F1 NIH_MGC_1c
gb_est2:BF971303	+	471.50	762.87	2.5e-33	658	BF971303 G02273150F1 NIH_MGC_84
gb_est2:BG714823	+	471.50	762.68	2.6e-33	674	BG714823 G02677146F1 NIH_MGC_9c
gb_est1:AV702285	+	471.50	762.49	2.6e-33	690	AV702285 AV702285 ADB Homo sapi
gb_est1:AV705808	+	471.50	762.48	2.6e-33	691	AV705808 AV705808 ADB Homo sapi
gb_est1:AL526209	+	471.50	762.43	2.6e-33	695	AL526209 AL526209 LTI_NFL003_NE
gb_est2:BG820179	+	471.50	762.39	2.7e-33	699	BG820179 G0278222F1 NCI_CGAP_E
gb_est1:BE790774	+	471.50	761.48	3.0e-33	783	BE790774 G01581956F1 NIH_MGC_7
gb_est1:AL526247	+	471.50	761.37	3.0e-33	794	AL526247 AL526247 LTI_NFL003_NE
gb_est1:AV702643	+	468.00	757.71	4.8e-33	683	AV702643 ADB Homo sapi
gb_est1:BE314909	+	468.00	756.01	4.7e-33	594	BE314909 G01140074F1 NIH_MGC_9
gb_est1:AL520931	+	468.00	756.51	5.7e-33	717	AL520931 AL520931 LTI_NFL004_NE
gb_est2:BE889313	+	468.00	755.92	6.1e-33	772	BE889313 G01513201F1 NIH_MGC_71
gb_est2:BG709503	+	468.00	755.54	6.4e-33	810	BG709503 G02674781F1 NIH_MGC_9c
gb_est2:BG707398	+	468.00	755.29	6.6e-33	835	BG707398 G02672810F1 NIH_MGC_9c
gb_est2:BG285825	+	468.00	754.30	7.5e-33	946	BG285825 G02380917F1 NIH_MGC_9c
gb_est2:BF038787	+	468.00	753.53	8.3e-33	1042	BF038787 G01462119F1 NIH_MGC_8
gb_est1:AI929703	+	464.50	752.59	9.4e-33	576	AI929703 G036f04.y1 Schneider 9
gb_est2:BF15659	+	464.50	750.65	1.2e-32	734	BF15659 G02676924F1 NIH_MGC_9c
gb_est2:BG714974	+	461.00	745.07	2.5e-32	726	BG714974 G02677263F1 NIH_MGC_9c
gb_est2:BF967675	+	461.00	744.36	2.7e-32	793	BF967675 G02287383F1 NIH_MGC_9c
gb_est2:BG707734	+	461.00	744.12	2.8e-32	817	BG707734 G02671255F1 NIH_MGC_9c
gb_est2:BG705843	+	457.00	737.68	6.5e-32	813	BG705843 G02669323F1 NIH_MGC_9c
gb_est2:BG705935	+	456.50	736.72	7.2e-32	828	BG705935 G026673343F1 NIH_MGC_9c

gb_est1:AI291270 + 455.00 738.14 6.0e-32 511 AI291270 qm16e06.x1 NCI_CGAP
gb_est1:AL024066 + 453.00 734.67 9.3e-32 526 AL024066 p8225a45 Beddington
gb_est1:BE266012 + 453.00 733.49 1.1e-31 610 BE266012 G01194731F1 NIH_MGC
gb_est1:AI291126 + 453.00 732.75 1.2e-31 569 AI291126 qm15f02.x1 NCI_CGAP
gb_est2:BF237433 + 453.00 732.27 1.3e-31 711 BF237433 G01842108F1 NIH_MGC

seq_name: gb_est1:AL022932

seq documentation block:
LOCUS AL022932 551 bp mRNA EST 29-DEC-1999
DEFINITION v8130b28 Beddington mouse dissected endoderm Mus musculus CDNA
clone 528-3B10 5', mRNA sequence.

ACCESSION AL022932

VERSION AL022932.1 GI:6645505

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 551)

AUTHORS Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and
Beddington,R.S.

TITLE Isolation of novel tissue-specific genes from cDNA libraries
representing the individual tissue constituents of the gastrulating

mouse embryo

JOURNAL Development 121 (8), 2479-2489 (1995)

MEDLINE 95401865

COMMENT Contact: Wiles,M., Lehrach,H. and Avner,P.
EEC Mouse Transcript Mapping Consortium

Genoscopy - CNS

2, rue Gaston Cremieux, 91000 Evry, France

Email: payner@pasteur.fr

clone available from Ressourcenzentrum, Heubnerweg 6, D-14059

Berlin, Germany. Web site http://www.rzpd.de

Seq primer: CCGGCCGGAATCCCGGT;

High quality sequence only submitted.

Vector: pSport1; site1: NotI; site2: SalI;

Cloned unidirectionally.

Dissected endoderm 7.5 days.

Average insert size: 1.2 kb (range: 0.2 - 2.kb).

Location/Qualifiers

1..551

Source

/organism="Mus musculus"

/strain="C57B16 x DBA"

/db_xref="taxon:10090"

/clone="528-3B10"

/clone.lib="Beddington mouse dissected endoderm"

/tissue_type="dissected endoderm"

/dev_stage="7.5 dpc"

/note="Vector: pSport1; Site1: NotI; Site2: SalI; Cloned

unidirectionally. High quality sequence only submitted."

Average insert size: 1.2 kb (range: 0.2 - 2.kb)"

BASE COUNT 153 a 124 c 191 g 83 t

ORIGIN

alignment_scores:

Quality: 646.00 Length: 128

Ratio: 5.168 Gaps: 0

Percent Similarity: 97.656 Percent Identity: 92.969

alignment_block:

US-09-327-750D-33 x AL022932

Align seg 1/1 to: AL022932 from: 1 to: 551

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17

|||||

142 ATGGAGTCCCAAGATCAAGCGGAAAAATCTCAACATGGAGATGACCA 191

|||||

17 sGlnLysLysGluGluLysGluLysProGlnAspThrLysArgG 34

|||||

192 TCAGAAAAGGAGGAGAGAGAAAAGCCACAGATACCATCAGAGAGG 241

34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
242 AGCAGCTGGCCCTGACCTCCGAGGCTGCCAAACTGTCGCCTAGA 291
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
292 GGAGGTGCGAGCGGTTCGGGTTCCGAGGCCATCGCTCACTATAGT 341
|||||
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
342 GGACCTGATGACAGAGGTTCGGGCTCCAGGAGGATGAGAGGAGA 391
|||||
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 100
|||||
392 ACGTACAGAGGTTCGGGCTGATGTGACACAGCTCATGGAGAGCTGAG 441
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
442 GAAGGAGCTGAGGACACAGCTCGGGGCTTAGCACTGACCGCCCTCA 491
|||||
117 sHisAspHisAspGluPheCysLeuMetPro 128
|||||
492 TCATGACCATCATGAGTTCCTGCTCATGCC 525
|||||

seq_name: gb_est1:AW536404

seq_documentation_block: 642 bp mRNA EST 31-AUG-2000
LOCUS AW536404
DEFINITION G0104A11-3 NTA Mouse E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0104A11 3', mRNA sequence.
ACCESSION AW536404
VERSION AW536404.1 GI:7178821
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
III, Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
CONTACT: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
Place: G0104 row: A column: 11
Seq primer: -21M13 Forward
High quality sequence stop: 642
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..642
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G0104A11"
/clone_lib="NTA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site1: SalI; Site2: NotI; Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an Oligo(dT) primer [NotI primer-adaptor
from GibcoBRL]
[5'-pGAGTACTGTTAGATCGGAGCGCGCTTTTTCCTTTT-3']
from 0.51ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by

ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Minoru S. H. Ko.

BASE COUNT 145 a 182 c 133 g 182 t
ORIGIN

alignment_scores:
Quality: 646.00 Length: 128
Ratio: 5.168 Gaps: 0
Percent Similarity: 97.656 Percent Identity: 92.969

alignment_block:
US-09-327-750D-33 x AW536404/rev ..

Align seg 1/1 to reverse of: AW536404 from: 1 to: 642

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17
|||||
606 ATGAGTGCCAAAGATCAAGCGGTGAAATCTCACTGGAGAAATGACCA 557
|||||
17 sGlnLysLysGluLysGluLysProGlnAspThrIleLysArgG 34
|||||
556 TCACAAAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
|||||
34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
506 AGCCAGCTGGCCCTGACCTCCGAGGCTGGCAAACTGTGCACCTAGA 457
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
456 GGAGTTCGACAGCGGTTCGGGTTCCGAGGCCATCGCTCACTATAGT 407
|||||
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
406 GGACCTGATGACAGAGGTTCGGGAGGCCCCAGGAGGAGTGAAGAGGAG 357
|||||
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 100
|||||
356 ACGTACAGAGGTTCGGGCTGATGTGACACAGCTCATGGAGAGCTGAGG 307
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
306 GAAGGAGCTGAGGACACAGCTCGGGGCTTAGCACTGACCGCCCTCA 257
|||||
117 sHisAspHisAspGluPheCysLeuMetPro 128
|||||
256 TCATGACCATCATGAGTTCCTGCTCATGCC 223
|||||

seq_name: gb_est2:BG064920

seq_documentation_block: 707 bp mRNA EST 26-JAN-2001
LOCUS BG064920
DEFINITION H3025D01-3 NIA Mouse 15K cDNA clone set Mus musculus cDNA clone
H3025D01 3', mRNA sequence.
ACCESSION BG064920
VERSION BG064920.1 GI:12547483
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 707)

AUTHORS Kargul G.J., Dudekula D.B., Qian Y., Lim M.K., Jaradat S.A., Tanaka T.S., Carter M.G. and Ko M.S.H.
TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other ESTs: H3025D01-5
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@gsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cdna/15k.html> for details.
 Plate: H3025 row: D column: 01
 Seq primer: -21M13 Forward
 High quality sequence stop: 707
 POLA=Yes

FEATURES Location/Qualifiers
 source
 1..707
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3025D01"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site1: SalI; Site2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryonic cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
 Mol Genet 7: 1967-1978."
 BASE COUNT 154 a 209 c 146 g 198 t
 ORIGIN

alignment_scores:
 Quality: 643.00 Length: 128
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 97.656 Percent Identity: 92.188
alignment_block:
 US-09-327-750d-33 x BG064920/rev ..
 Align seg 1/1 to reverse of: BG064920 from: 1 to: 707
 1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
 606 ATGGAGTCCAAAGATCAAGCGTGAAATAATCTCAACATGGAGAATGACCA 557
 17 scInLysLysGluGluLysProGlnAspThrLleLysArg 34
 556 TCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
 34 LuProValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
 506 AGCCAGCTGTGGCCCTGCCTCCGAGGCTGGCAAAACTGTGCACCTAGA 457
 51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyArgTr 67

|||||
 456 GGAGTCCGACGGTTCGGGTTCCGACGCCCATCGCTACATATAGATG 407
 67 PAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84
 406 GGACCTGATGACGAGGTTGGGGAGCCGCCGAGGAGGAGGAGGAGGAGG 357
 84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
 356 ACCTACACAGGTTTGGGGTGTGAGACAGCTCATGGAGAGCTGAGG 307
 101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
 306 GAAAGGACCTGAGCCACGACGCGTGGGGCGGTAGCACTGACCCGCTCA 257
 117 shisAspHisAspGluPheCysLeuMetPro 128
 256 TCATGACCACCATGATGAGTTTGCCTCATGCC 223
 seq_name: gb_est2:BG277659
 seq_documentation_block:
 LOCUS BG277659 575 bp mRNA EST 21-FEB-2001
 DEFINITION ux47c11.y1 Soares.NMMAX_maxillary.process Mus musculus cDNA clone
 INAMES:3513237 5' similar to TR:Q9RLJ2 Q9RLJ2 BEX1 PROTEIN. ; mRNA
 sequence.
 ACCESSION BG277659
 VERSION BG277659.1 GI:13073183
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 575)
 NC-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other ESTs: ux47c11.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaphs-r@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:1393813
 Seq primer: -40RP from Gibco
 High quality sequence stop: 462.
 Location/Qualifiers
 1..575
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:3513237"
 /clone_lib="Soares.NMMAX_maxillary.process"
 /tissue_type="maxillary process"
 /lab_host="DH10B (phage-resistant)"
 /note="Vector: pT7T3D-pac (Pharmacia) with a modified
 polylinker; Site1: NotI; Site2: EcoRI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAAGTGGGCGCGCGGCTTTTTTTTTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 156 a 136 c 201 g 81 t 1 others
 ORIGIN
alignment_scores:
 Quality: 633.00 Length: 128
 Ratio: 5.105 Gaps: 0
 Percent Similarity: 96.875 Percent Identity: 91.406
alignment_block:

US-09-327-750D-33 x BG277659

Align seg 1/1 to: BG277659 from: 1 to: 575

```

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
|||||
192 ATGGAGTCCAAAGATCAAGCGGTGAAAAATCTCAACATGGAGATGACCA 241
|||||
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34
|||||
242 TCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 291
|||||
34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
292 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 341
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
342 GGAGGTCCGAGCGGTTCGGGGTTCGGCAGCCCATCGCTCACTATAGATG 391
|||||
67 pAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
392 GCACCTCATGCAGAGGTTGGGAGGCCCGAGGAGGAGGAGGAGGAGGAG 441
|||||
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||
442 ACGTACAGAGGTTGGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 491
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
|||||
492 GAAAGGAGCTGTAGCCACAGCTCGCGGGGTTAGCAGCTGACCCGCTCA 541
|||||
117 sHisAspHisArgGluPheCysLeuMetPro 128
|||||
542 TCATGACCAACCATGATGATGTTGCTCATGCCCC 575
|||||

```

seq_name: gb_est2:BG870503

```

seq_documentation_block: 756 bp mRNA EST 29-MAY-2001
LOCUS BG870503 602791422F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922644 5',
DEFINITION mRNA sequence.
ACCESSION BG870503
VERSION BG870503.1 GI:14221043
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 756)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10842 row: 1 column: 05
High quality sequence stop: 756.

```

```

FEATURES
    source
        1..756
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4922644"
            /clone_lib="NCI_CGAP_SG2"
            /lab_host="DH10B (T1 phage-resistant)"
            /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:

```

NOTI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 207 a 164 c 230 g 155 t

ORIGIN

alignment_scores:
Quality: 633.00 Length: 129
Ratio: 5.064 Gaps: 1
Percent Similarity: 96.899 Percent Identity: 92.248

alignment_block:

US-09-327-750D-33 x BG870503

Align seg 1/1 to: BG870503 from: 1 to: 756

```

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
|||||
165 ATGGAGTCCAAAGATCAAGCGGTGAAAAATCTCAACATGGAGATGACCA 214
|||||
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34
|||||
215 TCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 264
|||||
34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
265 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 314
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
315 GGAGGTCCGAGCGGTTCGGGGTTCGGCAGCCCATCGCTCACTATAGATG 364
|||||
67 pAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluGlu 83
|||||
365 GCACCTCATGCAGAGGTTGGGAGGCCCGAGGAGGATGACGAGGAGGAG 414
|||||
84 AsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||
415 AACGTACAGAGGTTTGGGGGTGTGTGAGACAGCTCATGGAGAGCTGAG 464
|||||
100 gGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
|||||
465 GGAAGAGCAGCTGAGCCACAGCTTCGGGGGTTAGCAGCTGACCCGCTC 514
|||||
117 sHisAspHisArgGluPheCysLeuMetPro 128
|||||
515 ATCATGACCAACCATGATGATGTTTGCCTCATGCCCC 549
|||||

```

seq_name: gb_est1:BB654459

```

seq_documentation_block: 499 bp mRNA EST 06-SEP-2000
LOCUS BB654459 602791422F1 NIH_BMAP_MOB_N Mus musculus cDNA clone
DEFINITION UI-M-AJI-aha-f-10-0-UI-rl NIH_BMAP_MOB_N Mus musculus cDNA clone
UI-M-AJI-aha-f-10-0-UI 5', mRNA sequence.
ACCESSION BB654459
VERSION BB654459
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 499)
Normalizaton and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301-443 1706

```

Fax: 301 443 9890

Email: m8st@mail.nih.gov

cDNA Library Preparation: M. B. Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

source

1..499
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-AJ1-aha-f-10-0-UI"
 /clone_lib="NIH_BMAP_MOB_N"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT730-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_MOB_N library is a normalized library constructed from mouse olfactory bulbs. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dr track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories."

146 a 108 c 179 g 66 t

BASE COUNT
ORIGIN

alignment_scores:

Quality: 615.00 Length: 123
 Ratio: 5.125 Gaps: 0
 Percent Similarity: 97.561 Percent Identity: 92.683

alignment_block:

US-09-327-750D-33 x BE654459

Align seg 1/1 to: BE654459 from: 1 to: 499

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
 |||||
 131 ATGGAGTCAAGATCAAGCGGTGAAATCTCAATCATGGAGATGACCA 180
 17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34
 |||||
 181 TCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 230
 34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
 |||||
 231 AGCCAGCTGTGCCCTGACCTCCAGAGCTGGCAAAATCTGGACCTAGA 280
 51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
 |||||
 281 GGAGTCCAGCGGGTTCGGGGTTCGGCAGCCATCGCTACTATAGATG 330
 67 pasLeuMethIsArgValGlyGluProGlnGlyArgMetArgGluGlu 84
 |||||
 331 GGACCTGATGACAGGGTTCGGGAGCCCGGAGGAGGAGGAGGAGGAG 380
 84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
 |||||
 381 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTGAG 430
 101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
 |||||
 431 GAAAGGACAGCTGAGCCAGACCTGCGGGCGGTTAGCAGCTGACCCGCTCA 480
 117 shIsAspHisAspGlu 123
 |||||
 481 TCATGACCACCATGATGAG 499

seq_name: gb_est1:AW536974

seq_documentation_block:

LOCUS AW536974 577 bp mRNA EST 31-AUG-2000
 DEFINITION G0110H10-3 NIA Mouse E7.5 Embryonic Portion cDNA Library Mus
 musculus cDNA clone G0110H10 3', mRNA sequence.

ACCESSION AW536974

VERSION AW536974.1 GI:7179391

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 577)

REFERENCE

AUTHORS

Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,
 M.J., Pantano,S., Sano,Y., Piao,Y., NagaraJa,R., Doi,H., Wood,W.H.
 III, Becker,K.G. and Ko,M.S.H.

Genome-wide expression profiling of mid-gestation placenta and

embryo using a 15,000 mouse developmental cDNA microarray

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

20381348

MEDLINE

COMMENT

Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsn.grc.nia.nih.gov

Plate: G0110 row: H column: 10

Seq primer: -21M13 Forward

High quality sequence stop: 577

POLYA=Yes.

FEATURES

Location/Qualifiers

source

1..577
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G0110H10"
 /sex="unknown"
 /dev_stage="7.5dpc Embryo"
 /lab_host="DH10B"
 /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
 Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
 6 Embryo. The double-stranded cDNA was synthesized by
 Gibco's kit with an oligo(dT) primer [NotI primer-adaptor
 from GibcoBRL]
 [5'-pGACTAGTCTTAGATCGGAGCGCCCTTTTTTTTTTTT-3']
 from 0.5ug of mRNA. The double-stranded cDNAs were
 treated with T4 DNA polymerase and purified by
 ethanol-precipitation. The cDNAs were ligated to
 Lone-linker LI-Sal3 (include SalI sequence). The cDNAs
 were purified by phenol/chloroform and separated from
 free linkers by Centricon 100. Then, cDNAs were amplified
 by long-range high fidelity PCR using Takara's Ex Taq
 polymerase. Then, the cDNAs were purified by
 phenol/chloroform and by Centricon 100. The cDNAs were
 digested with SalI and NotI enzymes. Then, the cDNAs were
 size selected by Gibco's Size Fractionation Column. The
 cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by chemical method. The library was
 constructed by Xiaohong Wang and Minoru S. H. No."

BASE COUNT

ORIGIN

137 a 160 c 126 g 154 t

alignment_scores:

Quality: 599.00 Length: 118
 Ratio: 5.209 Gaps: 0
 Percent Similarity: 97.458 Percent Identity: 93.220

alignment_block:

US-09-327-750D-33 x AW536974/rev

Align seg 1/1 to reverse of: AW536974 from: 1 to: 577

```

11 LeuAsnMetGluAsnAspHisGlnLysLysGluLysGluLysGluLysPr 27
576 CTCACATGGAGTAATGACCATCAGAAAAGAGGAGGAGGAAGAAAGGCC 527
27 oGlnAspThrIleLysArgGluProValAlaProThrPheGluAla 44
526 ACAAGATACCATCAGAAAGGAGCCAGCTGTGGCCCTGACCTCCGAGGCTG 477
44 LysAsnGlnCysAlaProArgGlyGlyArgArgPheArgValArgGln 60
476 GCAAAACCTGTCCACTAGAGAGGTCCAGCGGTTCGGGTTCGCGAG 427
61 ProIleSerHisTyrArgTrpAspLeuMetHisArgValGlyGluProG1 77
426 CCCATCGCTCACTATAGATGGACCTGATGACAGAGGTTGGGAGGCCCA 377
77 nGlyArgMetArgGluGlnAsnValGlnArgPheGlyGluAspMetArg 94
376 GGGAGGATGAGAGAGGAGACGATACAGAGTTTGGGGGTGTGTGAGAC 327
94 InLeuMetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAla 110
326 AGCTCATGAGAAGCTGAGGGAAGGACGCTGAGCCACAGCTGCGGGCG 277
111 ValSerThrAspProHisHisAspHisHisAspGluPheCysLeuMe 127
276 GTTAGCAGTGCACCCGCTCATCATGACACCATGATGAGTTTGCCTCAT 227
127 tPro 128
226 GCCC 223

```

seq_name: gb_est2.BF607762

seq_documentation_block: 785 bp mRNA EST 01-APR-2001
LOCUS MY1_000704 Mouse 9-day fetus cDNA library ICRP522 Mus musculus
DEFINITION cDNA clone ICRP522A1947 5', mRNA sequence.
ACCESSION BF607762
VERSION BF607762.1 GI:13504254
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B. G.,
Lehrach, H. and O'Brien, J.
TITLE Detection of a high number of novel genes in a 9-day mouse embryo
JOURNAL cDNA library normalised by oligonucleotide fingerprinting
COMMENT Unpublished (2001)
Contact: Hennig S
Laboraty 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
EST's are made from clones being representatives of clone clusters.
clone clusters were calculated from oligonucleotide fingerprints.
PCR Primers
FORWARD: 5'-GAGCTATTCAGAGTAGTGA-3'
BACKWARD: 5'-TAATACGACTACTATAGGG-3'
Seq primer: 5'-ATTAGTGACACTAG-3'
High quality sequence stop: 785.
Location/Qualifiers
1..785
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="ICRP522A1947"
/clone_lib="Mouse 9-day fetus cDNA library ICRP522"

FEATURES
source

/tissue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E.coli, XL1 blue"
/note="vector: pSVSPORT1; Site_1: NotI; Site_2: SalI;
Library preparation by oligo-dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."
BASE COUNT 190 a 188 c 233 g 153 t 21 others
ORIGIN

alignment_scores:
Quality: 554.00 Length: 128
Ratio: 4.617 Gaps: 4
Percent similarity: 93.750 Percent identity: 87.500
alignment_block:
US-09-327-750D-33 x BF607762

Align seg 1/1 to: BF607762 from: 1 to: 785

```

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
259 ATGGAGTCCAAAGATCAAGCGCTGAAAAATCTCAACATGGAGATGACCA 308
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34
309 TCAGAAAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 358
34 LuProValValAlaProThrPheGluAlaGlyLysAsnGlnAlaProArg 50
359 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGCCAAAACCTGTCACCTAGA 408
51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyrArgT 67
409 GGAGGTGCGAGCGGTTCGGGTTCGGCAGCCATCGCTCCTACTAGATG 458
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84
459 GGACCTGTGTCAGAGGAGGTTCGGGAGAGAGAGAGAGAGAGAGAGAG 508
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuAr 100
509 ACGTACAGAGTTTGGGGGTGATGTGAGACAGCTCATGGGGAGCTGAG 558
100 gGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp.ProPro 116
559 GGAAGGACACCTGAACACAGAGCTGCGGGCGGTAGCAGCTGACCCGCT 608
117 HisHisAspHisHisAspGluPhe 124
609 CATTATGACCAACCATGGATGAGTTT 634

```

seq_name: gb_est2.BG228077

seq_documentation_block:
LOCUS BG228077 587 bp mRNA EST 08-FEB-2001
DEFINITION ux47c11.x1 Soares_NMMAX_maxillary_process Mus musculus cDNA clone
IMAGE:3513237 3' similar to TR:Q9RIJ2 Q9RIJ2 BEX1 PROTEIN. ; mRNA
sequence.
ACCESSION BG228077
VERSION BG228077.1 GI:12715592
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 587)
TITLE NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

456 ACGTACAGAGTCTGGGGTGATGTGAGACAGCTCATGGAGAACTCAGG 505

101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPtoHI 117

506 GAAAGGAGCTTATGACACAGCTCGGGGGGTAGCACTGACNCGCCTCA 555

117 sHisAspHisAspGluPheCysLeuMetPro 128:

556 TCATGACTACCATGATTAGTTTTCCTCATGCCC 589

seq_name: gb_est1.BE291071

seq_documentation_block:

LOCUS BE291071 583 bp mRNA EST 13-JUL-2000

DEFINITION 601086311F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500522 5'

RNA sequence.

ACCESSION BE291071

VERSION BE291071.1 GI:9172545

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

REFERENCE 1 (bases 1 to 583)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14W8559 row: e column: 03

High quality sequence start: 5

High quality sequence stop: 503.

Location/Qualifiers

FEATURES

source

1..583

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="3500522"

/clone_lib="NCI_CGAP_Mam6"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Jeffrey Green, M.D., NIH

165 a 136 c 201 g 81 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 505.50 Length: 120

Ratio: 4.554 Gaps: 2

Percent Similarity: 92.500 Percent Identity: 83.333

alignment_block:

US-09-327-750D-33 x BE291071

Align seg 1/1 to: BE291071 from: 1 to: 583

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHI 17

|||||

208 ATGGAGTCCAAAGATCAAGCGGTGAATAATCTCAACATGGAGATGACCA 257

|||||

17 sGlnLysLysGluLysGluLysProGlnAspThrIleLysArgG 34

258 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 307

34 LuProValValAlaProThrPheGluAlaGlyAsnCysAlaProArg 50

308 ACCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 357

51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67

358 GGAGGTGCAGAGCGGTTCGGGGTTCGGCAGCCCATCGCTCACTATAGTG 407

67 PASLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGua 84

408 GGACCTCATCAGAGGTTGGGGAGCCCGCCAGTGAAGGATGAGAGAGGAGA 457

84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100

458 ACGTACAGAGGTTAGGGGGTGTGTGAGACAGCTCATGGAGAAGCTGAGG 507

101 GluArgGlnLeuSerHis...

508 GAAAGGCTGTGAGCCACAGACATCCGAGGCGGTTAGACAATGACCCGA 557

116 roHisHis 118

558 CTCATCAT 565

seq_name: gb_gss:A2936393

seq_documentation_block:

LOCUS A2936393 637 bp DNA GSS 26-APR-2001

DEFINITION 2M0193105F Mouse 10kb plasmid UUGC2M library Mus musculus genomic

clone UUGC2M0193105 F, DNA sequence.

ACCESSION A2936393

VERSION A2936393.1 GI:13794974

KEYWORDS GSS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 637)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0193 row: L column: 05

Seq primer: CGTTGTAAACGACGCGCAGT

Class: plasmid ends

High quality sequence stop: 637.

Location/Qualifiers

1..637

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone_image="UUGC2M0193105"

/clone_lib="Mouse 10kb plasmid UUGC2M library"

/sex="Female"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (female) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g147321149b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 131 a 137 c 132 g 176 t 1 others
ORIGIN

alignment_scores:
Quality: 493.50 Length: 117
Ratio: 4.569 Gaps: 1
Percent Similarity: 92.308 Percent Identity: 84.615

alignment_block:

US-09-327-750D-33 x A2936393/rev ..

Align seg 1/1 to reverse of: A2936393 from: 1 to: 637

```

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16
382 ATGGAGTCCAAAGTGGACAAAGCGGTGAAATAATCTCAACATGGAGAATGA 333
16 pHisGlnLysLysGluGluGluLysGluGluProGlnAspThrIleLysA 33
332 CCATCAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 283
33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
282 GGGATCCGATGTGGCCCTTCCTTCGAGAGCTGGAGACTACTACGTGCCT 233
50 ArgGlyClyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
232 AGAGGAGTCCGAGGCGGTTCGGGTTCGGCAGCCATCGTCACACTACAG 183
66 gTTPAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
182 ATGGAGCTGATCATAGAGGTGGGAGGCCCGGAGGAGGAGGAGGAGGAGG 133
83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
132 AGAACGTACAGAGGTTTGGGGATGATGTGAGACAGCTCATGGAGAGCTN 83
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
82 GG.GAAAGGCGCTGAGCCACACACCTTCGGCGCGGTTCACACTGACCCGCC 34
116 o 116
33 T 33

```

seq_name: gb_est2:BG669326

seq_documentation_block:
LOCUS BG669326 412 bp mRNA EST 30-APR-2001
DEFINITION DRNAEB11 Rat DRG library Rattus norvegicus cDNA clone DRNAEB11
5', mRNA sequence.
ACCESSION BG669326
VERSION BG669326.1 GI:13891248
KEYWORDS EST
SOURCE Norway rat.
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 412)
Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.
, Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,W.G., Chen,Z. and
Zhang,X.

TITLE
Distinct gene expression profiles of rat dorsal root ganglion
induced by peripheral nerve axotomy

JOURNAL
COMMENT

Unpublished (2001)

Contact: Zhang Xu

Laboratory of Sensory System

Institute of Neuroscience

320 Yue Yang Road, Shanghai 200031, P.R.China

Tel: 86-21-64748700-121

Fax: 86-21-64713446

Email: xu.zhang@ion.ac.cn

This clone is also available at Chinese National Human Genome
Center at Shanghai, 351 Guo Shouling Road, Zhangjiang Hi-Tech Park,
Pudong New Area, P.R.China. Please contact with Zhang Xu
(xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@chgc.sh.cn)

PCR Primers

BACKWARD: T3

Seq primer: T3

POLYA-No.

Location/Qualifiers

1..412

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="DRNAEB11"

/clone_lib="Rat DRG Library"

/sex="male"

/tissue_type="dorsal root ganglion"

/dev_stage="adult"

BASE COUNT 124 a 83 c 130 g 75 t

ORIGIN

alignment_scores:

Quality: 489.50 Length: 126
Ratio: 4.450 Gaps: 2
Percent Similarity: 87.302 Percent Identity: 76.984

alignment_block:

US-09-327-750D-33 x BG669326 ..

Align seg 1/1 to: BG669326 from: 1 to: 412

```

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16
37 ATGGAGTCCAAAGTGGACAAAGCGGTGAAATAATCTCAACATGGAGAATGA 86
16 pHisGlnLysLysGluGluGluLysGluGluProGlnAspThrIleLysA 33
87 CCATCAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
137 GGGAGCGGTTGTAGCCCTTCCTTCGAGAGCTGGAGACTACTACGTGCCT 186
50 ArgGlyClyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
187 ATAGAACTCCAGCGGTTCCCGGTTTGGCAGCCCATTTGCCACTATAT 236
66 g.TTPAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 82
237 GATGGACCTGATCGCAAGAGTTGGAGAGCCCAAGCAGGAGGATGAGAGAA 286
83 GluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLe 99
287 GAGATGTGCAAGGTTTGGGAGGATATGAGACAGCTCATGTGAATTT 336

```

100

```

/db_xref="taxon:10090"
/clone="G0106H08"
/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH108"
/notes="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an oligo(dT) primer [NotI primer-adaptor
from GibcoBRL]
[5'-pGACTAGTCTAGATCGGCGCCCTTTT-3']
from 0.51ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker LL-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH108 E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Minoru S. H. Ko."
BASE COUNT      151 a  151 c  117 g  130 t
ORIGIN

```

```

alignment_scores:
  Quality: 472.00      Length: 97
  Ratio: 5.187        Gaps: 0
  Percent Similarity: 93.814  Percent Identity: 88.660
alignment_block:
US-09-327-750D-33 x AW536634/rev
Align seg 1/1 to reverse of: AW536634 from: 1 to: 549
32 LysArgGluProValAlaProThrPheGluAlaGlyLysAsnCysAl 48
|||||:|||||:|||||:|||||:|||||:|||||:
541 AAAGGGATCCGATTGTGGCCCTTCGAGCTGGAGACTACTACGT 492
48 aProArgGlyArgArgPheArgValArgGlnProIleSerHist 65
|||||:|||||:|||||:|||||:|||||:|||||:
491 GCCTAGAGGAGGTCCAGCGGTTCCGGTTCCGACCCATCGTCACT 442
65 yArgTrpAspLeuMethHisArgValGlyGluProGlnGlyArgMetArg 81
|||||:|||||:|||||:|||||:|||||:|||||:
441 ACAGATGGGACCTGATGCTAGGTTGGGAGGCCGCCAGGAAGGATGAGA 392
82 GluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLy 98
|||||:|||||:|||||:|||||:|||||:|||||:
391 GAGGAGACGTACAGAGGTTTGGGGATGATGTGAGACAGCTCATGGAGAA 342
98 sLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp 115
|||||:|||||:|||||:|||||:|||||:|||||:
341 GCTGAGGGAAGGACAGCTGAGCCACAGCTCCGGCGGTTAGCACCTGACC 292
115 rProHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:
291 CGCCTCATCATGACCACCATGATGATTTTGCCTTATGCC 251

```

/cgn2_6/ptodata/2/ina/5B_COMB.seq	US-08-452-0763-1	-	74.50	120.90	43.90	2841
/cgn2_6/ptodata/2/ina/6B_COMB.seq	US-09-231-063-1	-	74.50	120.90	43.90	2841
/cgn2_6/ptodata/2/ina/5B_COMB.seq	US-08-403-852D-1	+	74.50	114.40	101.08	5392
/cgn2_6/ptodata/2/ina/6A_COMB.seq	US-08-510-646B-1	+	74.50	114.40	101.08	5392
/cgn2_6/ptodata/2/ina/6B_COMB.seq	US-09-231-818-1	+	74.50	114.40	101.08	5392

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-381-881-1

```

seq_documentation_block:
; Sequence 1, Application US/08381881
; Patent No. 593064
;
; GENERAL INFORMATION:
;
; APPLICANT: SAVELKOU, PAUL H. M.
; APPLICANT: GAASTRA, WILLEM
; TITLE OF INVENTION: BORDETELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381.881

```

Sequence	Strd Orig	ZScore	EScore Len	Documentation
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-381-881-1 +		93.50	166.92	1315
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-281-221-1 +		93.50	166.92	1315
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-07-951-715A-6 +		93.00	155.63	3624
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-459-448A-6 +		93.00	155.63	3624
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-459-459A-6 +		93.00	155.63	3624
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-439-504B-6 +		93.00	155.63	3624
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-459-444-6 +		93.00	155.63	3624
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-053-549-7 +		93.00	155.63	3624
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-053-549-7 +		91.00	142.54	8854
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-506-453C-6 +		82.50	152.65	607
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-988-416-994 -		82.00	149.79	729
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-426-599B-3 +		81.00	149.35	624
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-147-777-3 +		80.50	110.90	24979
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-452-872-3 +		80.50	110.90	24979
/cgn2_6/ptodata/2/ina/PTCUS.COMB.seq:PTC-US93-03985-3		80.50	110.90	158.26
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-434-288-10 +		80.00	137.79	5.03
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-030-995-1 +		79.00	137.39	5.30
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-117-927-4 +		78.50	131.71	10.97
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-931-999-4		77.50	118.14	62.53
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-125-468-1 -		77.50	103.01	435.42
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-474-933-1 -		77.50	103.01	435.42
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-998-416-141 -		77.00	140.56	3.23
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-047-041A-24 -		77.00	133.96	8.23
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-285-440-11 +		76.50	130.76	12.40
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-630-345-11 +		76.50	130.76	12.40
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-285-440-13 +		76.50	127.81	18.09
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-630-349-13 -		76.50	127.81	18.09
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-306-691B-23 +		76.50	127.06	19.92
/cgn2_6/ptodata/2/ina/PTCUS.COMB.seq:PTC-US93-06251-78 +		76.50	127.06	19.92
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-056-200-93 +		76.50	112.62	127.02
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-800-644-93 +		76.50	112.62	127.02
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-036-987A-1 -		76.00	90.02	2.3e+03
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-370-700-1 -		76.00	90.02	2.3e+03
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-717-294-42 +		75.50	118.36	60.85
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-103-840A-2 +		75.50	48.36	3.5e+05
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-381-881-3 +		75.00	137.29	5.37
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-281-221-3 +		75.00	137.29	5.37
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-1 -		75.00	120.54	46.01
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-4 +		75.00	120.54	46.01
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-4 +		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-1 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-4 +		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-1 -		75.00	109.04	20.00
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 -		75.00	109.04	

```
alignment_scores:
  Quality: 93.50      Length: 96
  Ratio: 1.948      Gaps: 4
  Percent Similarity: 50.000      Percent identity: 33.333

alignment_block:
US-09-327-750D-33 x US-08-381-881-1 ..
```

```

Align seg 1/1 to: US-08-381-881-1 from: 1 to: 1315
43 AlaglyLysAsnCysAlaProArgGlyGlyArgGlyArgPheArgValar 59
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
542 GCAAGCCAAACAGCTTCTCTGGCGCGCGCTCGCGCGCTCGCGCTCG 591
59 gGlnProIle.....SerHisTyArgTrp..... 67
592 CCGCCATCGGAGAGCGCACCATTTGTCATTACCGGCGCAGCATCCGAC 641
68 ..AspLeuMetHisArgValGlyGluPro.....GlnGlyArgMet 80
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
642 CAGACCTGCAGTCAGGACCGCGCGCGCTTACATCAAGTGTGCTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetG1 97
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
692 CCT.....GCCACGATCTCCAAGAGCGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
721 AGACGCGCGCGAGTGGCGGCGCGCGCTCGCTTCATATCAAGTGAAG 770
114 spProProHisHisAspHisAspGluPheCysLeu 126
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
771 GACTGCCCGACCGCTCAACACTCTCAAGCTGTACTT 808

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-281-221-1

seq_documentation_block:
; Sequence 1, Application US/09281221
; Patent No. 6284256
; GENERAL INFORMATION:
; APPLICANT: SAVALKOU, PAUL H., M.
; TITLE OF INVENTION: BORDETELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/281,221
; FILING DATE: 30-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/381,881
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TTP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; TELEX: 44-0704 "STEVENS"
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```

```

ORIGINAL SOURCE:
ORGANISM: Bordetella bronchiseptica
STRAIN: 401
IMMEDIATE SOURCE:
CLONE: E coli PC2495(pIVB3-420)
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..539
FEATURE:
NAME/KEY: CDS
LOCATION: 540..1142
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1143..1315
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-281-221-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.948 Gaps: 4
Percent Similarity: 50.000 Percent Identity: 33.333

alignment_block:
US-09-327-750D-33 x US-09-281-221-1 ..
Align seg 1/1 to: US-09-281-221-1 from: 1 to: 1315
43 AlaglyLysAsnCysAlaProArgGlyGlyArgGlyArgPheArgValar 59
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
542 GCAAGCCAAACAGCTTCTCTGGCGCGCGCTCGCGCGCTCGCGCTCG 591
59 gGlnProIle.....SerHisTyArgTrp..... 67
592 CCGCCATCGGAGAGCGCACCATTTGTCATTACCGGCGCAGCATCCGAC 641
68 ..AspLeuMetHisArgValGlyGluPro.....GlnGlyArgMet 80
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
642 CAGACCTGCAGTCAGGACCGCGCGCGCTTACATCAAGTGTGCTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetG1 97
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
692 CCT.....GCCACGATCTCCAAGAGCGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
721 AGACGCGCGCGAGTGGCGGCGCGCGCTCGCTTCATATCAAGTGAAG 770
114 spProProHisHisAspHisAspGluPheCysLeu 126
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
771 GACTGCCCGACCGCTCAACACTCTCAAGCTGTACTT 808

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-07-715A-6

seq_documentation_block:
; Sequence 6, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Kozziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rochstein, Steven J.
; APPLICANT: Bowman, John L.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.

```

```
APPLICANT: Suttle, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA: US/07/951,715A
FILING DATE: 25-SEP-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Sprull, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
OTHER INFORMATION: Optimized cryIB"
OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-07-951-715A-6

alignment_scores:
Quality: 93.00 Length: 116
Ratio: 1.550 Gaps: 4
Percent Similarity: 51.724 Percent Identity: 26.724

alignment_block:
US-09-327-750D-33 x US-07-951-715A-6

Align seg 1/1 to: US-07-951-715A-6 from: 1 to: 3624

14 GluAsnAspHisGlnLysLysGluGluLysGluGluLysProGlnAspHis 30
|||||
1026 GAACACAGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1075
30 rleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
|||||
1076 TGCCTGTCGCGCGCGCGAGCTGTACCGCAGCAGCAGCAGCAGCAGCGTG 1125
47 ysAla.....ProArgGlyGlyArgArg 55
|||||
1126 CTGCTGTGGGCGCATCTACCTGGAGCCCATCCAGCGGTGCCACCGCGTG 1175
56 PheArgValArgGlnProIleSerHisTyrArgTrpAspLeuMethHisar 72
|||||
1176 CTTCACTTCAACCAACCCCAAGAACATCAGCGA.....CCGCGGCACCG 1219
```

```
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
|||||
1220 CCACTACAGCCA.....GCCCTACAGAGAGAGCCCGCCTGCAGCTGAAG 1263
89 lyGluAspMetArgGlnLeuMetGlnLysLeuArgGluArgGlnLeuSer 105
|||||
1264 GACACGAGACCGAGCTGCCCGCAGACCGACCGAGCGCCCA..... 1307
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHis 121
|||||
1308 .....CTACGAGAGTACAGCCACCGCCTGACGACCATCGGCATCAT 1349
seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-459-448A-6

seq_documentation_block:
; Sequence 6, Application US/08459448A
; Patent No. 5853336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttle, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,448A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: Optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
; US-08-459-448A-6
;
; alignment_scores:
; Quality: 93.00 Length: 116
; Ratio: 1.550 Gaps: 4
; Percent Similarity: 51.724 Percent Identity: 26.724
;
; alignment_block:
; US-09-327-750D-33 x US-08-459-448A-6
;
; Align seg 1/1 to: US-08-459-448A-6 from: 1 to: 3624
;
; 14 GluAsnAspHisGlnLysGluLysGluLysGluLysProGlnAspTh 30
; 1026 GAACACGAGCAGCCAGCGCGCCACCAACACGAGCATCAACCCGCGTACCC 1075
;
; 30 rleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
; 1076 TGCCTTTCGCCAGCGCGAGTGTACCGCAGCGAGCATCGCGCGGTG 1125
;
; 47 ysAla.....ProArgGlyGlyArgArg 55
; 1126 CTGCTGTGGGCGATCTACCTGGAGCCCATCCAGCGGTGCCACCGTGGC 1175
;
; 56 PheArgValArgGlnProIleSerHisTyArgTrpAspLeuMetHisAr 72
; 1176 CTTCACTTCACCAACCCCGACCAACATCAGCA.....CGCGGCGACCG 1219
;
; 72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
; 1220 CCAACTACAGCCA.....GCCCTAGAGAGCGCGCGCTGCAGCTGAAG 1263
;
; 89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGluLeuSer 105
; 1264 GACAGCAGACCGAGCTGCCCGCCGAGACCCAGCCAGCGCCCA..... 1307
;
; 106 HisSerLeuArgAlaValSerThrAspProProHisHisAspHisHis 121
; 1308 .....CTACGAGAGCTACAGCAGCGCGCTGAGCCACATCGGCATCAT 1349
;
; seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-459-595A-6
;
; seq_documentation_block:
; ; Sequence 6, Application US/08459595A
; ; Patent No. 6018104
; ; GENERAL INFORMATION:
; ; APPLICANT: Koziel, Michael G.
; ; APPLICANT: Desai, Nalini M.
; ; APPLICANT: Lewis, Kelly S.
; ; APPLICANT: Kramer, Vance C.
; ; APPLICANT: Warren, Gregory W.
; ; APPLICANT: Evola, Stephen V.
; ; APPLICANT: Crossland, Lyle D.
; ; APPLICANT: Wright, Martha S.
; ; APPLICANT: Merlin, Ellis J.
; ; APPLICANT: Launis, Karen L.
; ; APPLICANT: Rothstein, Steven J.
; ; APPLICANT: Bowman, Cindy G.
; ; APPLICANT: Dawson, John L.
; ; APPLICANT: Dunder, Erik M.
; ; APPLICANT: Pace, Gary M.
; ; APPLICANT: Suttie, Janet L.

```

```

;
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018104artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,595A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: Optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
; US-08-459-595A-6
;
; alignment_scores:
; Quality: 93.00 Length: 116
; Ratio: 1.550 Gaps: 4
; Percent Similarity: 51.724 Percent Identity: 26.724
;
; alignment_block:
; US-09-327-750D-33 x US-08-459-595A-6
;
; Align seg 1/1 to: US-08-459-595A-6 from: 1 to: 3624
;
; 14 GluAsnAspHisGlnLysGluLysGluLysGluLysProGlnAspTh 30
; 1026 GAACACGAGCAGCCAGCGCGCCACCAACACGAGCATCAACCCGCGTACCC 1075
;
; 30 rleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
; 1076 TGCCTTTCGCCAGCGCGAGTGTACCGCAGCGAGCATCGCGCGGTG 1125
;
; 47 ysAla.....ProArgGlyGlyArgArg 55
; 1126 CTGCTGTGGGCGATCTACCTGGAGCCCATCCAGCGGTGCCACCGTGGC 1175

```



```

;
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,444A
; FILING DATE: 02-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized civB"
; /note= "Disclosed in Figure 6."
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-459-444-6

```

```

alignment_scores:
  Quality: 93.00      Length: 116
  Ratio: 1.550       Gaps: 4
  Percent Similarity: 51.724      Percent Identity: 26.724

```

```
alignment_block:
US-09-327-750D-33 x US-08-459-444-6

```

```
Align seg 1/1 to: US-08-459-444-6 from: 1 to: 3624
```

```

14  GluAsnAspHisGlnLysLysGluGluLysGluLysProGlnAspTh 30
|||||
1026 GAACACAGACCCAGCGGCCACACACACAGCAGCATCACCCGCGGCC 1075
30  rIleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
|||||
1076 TGCCTTCGCGAGCGCGAGCTGTACCGACGAGAGTACCGCGCGTG 1125
47  ySaLa.....ProArgGlyClyArgArg 55
|||||
1126 CTGTGTGGGCATCTACCTGGAGCCCATCCAGCGGTGCCACCGGTGCG 1175
56  PheArgValArgGlnProIleSerHisTyrArgTrpAspLeuMethHis 72
|||||
1176 CTTCAACTTACCAACCCCGGACGAGACATCAGCGA.....CGCGGCGACCG 1219

```

```

72  gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
|||||
1220 CCAACTACAGCA.....GCCCTACGAGAGCCCGCGCTGCAGCTGAAG 1263
89  lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSer 105
|||||
1264 GACAGCGAGACCGAGCTGCCCGCCGAGACACCGAGCGGCCCA..... 1307
106  HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
|||||
1308 .....CTACGAGAGCTACAGCCCGCTGAGCCACATCGGCATCAT 1349
seq_name: /cgn2.5/ptodata/2/lna/6A_COMB.seq:US-09-053-549-7

```

```

seq_documentation_block:
; Sequence 7, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521lel Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-Apr-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized civB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-09-053-549-7

```

```

alignment_scores:
  Quality: 93.00      Length: 116
  Ratio: 1.550       Gaps: 4
  Percent Similarity: 51.724      Percent Identity: 26.724

```

```
alignment_block:
US-09-327-750D-33 x US-09-053-549-7

```

```
Align seg 1/1 to: US-09-053-549-7 from: 1 to: 3624
```

```
14 GluAsnAspHisGlnLysLysGluGluLysGluGluLysProGlnAspTh 30
|||||
1026 GAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1075
30 rIleLysArgGluProValAlaProThrPheGluAlaGlyLysAsnC 47
|||||
1076 TCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1125
47 ysAla.....ProArgGlyGlyArgArg 55
|||||
1126 CTGCTGTGGGCACTACCTGGAGCCCATCCACGCGCGCGCGCGCGCG 1175
56 PheArgValArgGlnProIleSerHisTyArgTrpAspLeuMetHisAr 72
|||||
1176 CTCAACTTCAACACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1219
72 gValGlyGluProGlnGlyArgMetArgGluGluGluGluGluGluGlu 89
|||||
1220 CCAACTACAGCCA.....GCCCTAGCAGAGCGCGCGCGCGCGCGCG 1263
89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluGluGluGluSer 105
|||||
1264 GACAGCGAGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1307
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
|||||
1308 .....CTACGAGAGCTACAGCCAGCGCGCGCGCGCGCGCGCGCAT 1349
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-09-053-549-1
seq_documentation_block:
; Sequence 1, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalin
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..3694
; OTHER INFORMATION: /product- "hyFLIB protein"
```

```
; NAME/KEY: misc_feature
; LOCATION: 8854
; OTHER INFORMATION: /note- "Sequence of pCIB5520
; Patent No. 6121521
; OTHER INFORMATION: containing coding sequence for hyFLIB protein"
US-09-053-549-1

alignment_scores:
Quality: 91.00 Length: 116
Ratio: 1.542 Gaps: 4
Percent Similarity: 50.862 Percent Identity: 27.586
alignment_block:
US-09-327-750D-3 x US-09-053-549-1
Align seg 1/1 to: US-09-053-549-1 from: 1 to: 8854
14 GluAsnAspHisGlnLysLysGluGluLysGluGluLysProGlnAspTh 30
|||||
1099 GAACACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1148
30 rIleLysArgGluProValAlaProThrPheGluAlaGlyLysAsnC 47
|||||
1149 TCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1198
47 ysAla.....ProArgGlyGlyArgArg 55
|||||
1199 CTGCTGTGGGCACTACCTGGAGCCCATCCACGCGCGCGCGCGCGCG 1248
56 PheArgValArgGlnProIleSerHisTyArgTrpAspLeuMetHisAr 72
|||||
1249 CTCAACTTCAACACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1292
72 gValGlyGluProGlnGlyArgMetArgGluGluGluGluGluGluGlu 89
|||||
1293 CCAACTACAGCCA.....GCCCTAGCAGAGCGCGCGCGCGCGCGCG 1336
89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluGluGluGluSer 105
|||||
1337 GACAGCGAGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
|||||
1381 .....CTACGAGAGCTACAGCCAGCGCGCGCGCGCGCGCGCGCAT 1422
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-08-553C-6
seq_documentation_block:
; Sequence 6, Application US/08506553C
; Patent No. 6120989
; GENERAL INFORMATION:
; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Schoeneborn, Han-H.;
; APPLICANT: Plachter, Bodo; and Jahn, Gerhard
; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/506,553C
; FILING DATE: 07/25/95
; CLASSIFICATION: 433
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LEADER 202-PFF/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-506-553C-6

alignment_scores:
  Quality: 82.50      Length: 117
  Ratio: 1.650       Gaps: 4
  Percent Similarity: 42.735      Percent Identity: 25.641

alignment_block:
US-09-327-750D-33 x US-08-506-553C-6
..
Align seg 1/1 to: US-08-506-553C-6 from: 1 to: 607

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyAr 66
|||||
107 CGGTCTGGAGGAGTGGCGCGGTGGCGCAAGAA...GCACGACCGCG 153
|||||
66 tTtPAsPLeuMethHisArgValGluProGln..... 77
|||||
154 GTGGCGGCGGTGGTTCGGTACGCGGAATGAGTAGCGGTGGCGCGGC 203
77 .....

204 GGTGATCATGACCAGCGTCTTCTCTCAAGAGAAATACGACGACAA 253
78 .....GlyArgMetArgGlu 83
254 GATCACAGCTACTGACTGCTCAAGAGTGATCGCGCGCGGAGGAG 303
|||||
83 uAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 100
|||||
304 GAGGAGCGCGGTTGGATCGCAACTCCGCAATTACTTCAACGACGCG 353
100 tG.....GluArgGlnLeuSer..... 105
354 AAAGAGGAGCGGACGCGGAGGATCTGTACCTTCGATTCGTCCTAA 403
106 HisSerLeuArgAlaValSerThrAspProProHisHisAspHisAs 122
|||||
404 CACCAAGAAGCAAAAGTGGCGGAGATCTCTGAAGAGCAGCGGCGCAT 453
122 p 122
454 A 454

```

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-994

```

```

seq_documentation_block:
; Sequence 994, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippesen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jürgen

```

```

; APPLICANT: knechtie, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 994:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG16100P
; US-08-998-416-994

alignment_scores:
  Quality: 82.00      Length: 112
  Ratio: 1.302       Gaps: 8
  Percent Similarity: 56.250      Percent Identity: 31.250

alignment_block:
US-09-327-750D-33 x US-08-998-416-994/rev ..
..
Align seg 1/1 to reverse of: US-08-998-416-994 from: 1 to: 729

16 AspHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIle 32
|||||
670 AACACAGAAAGCTCGACAGCAACAGTGTGCGCCAGTT..... 632
|||||
32 sArgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlap 49
|||||
631 .....GAACCATATATATACCGTTTCCAC.....GCCA 601
49 tArG.GlyGlyArgArgPhe.ArgValArgGlnProIleSerHisT 65
|||||
600 AACCGAGCGGACAGAAACCGCTTTGAAATATACCGTCCGAGGAGTAC 552
65 yTArgTtPAsPLeuMethHisArgValGluGluProGlnGlyArgMetArg 81
|||||
551 ..AGGATGAACGTCGCGCATCTC.....CCCCAG...CGACTGCGT 516
82 GluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeu..... 95
|||||
515 GATCGAGTACACCGTACCGATGATGATGATGATGATGATGATGATGAT 466

```



```

alignment_scores:
  Quality: 80.50      Length: 72
  Ratio: 1.872      Gaps: 1
  Percent Similarity: 59.722      Percent Identity: 31.944

alignment_block:
  US-09-327-750D-33 x US-08-147-777-3
  Align seg 1/1 to: US-08-147-777-3 from: 1 to: 24979

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyAr 66
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
5088 CGCGGAGGCGGCGGAGCTACGCTCAGAGGCTACGCGGCGGCGGAGCG 5137

66 gTTPAsPLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluG 83
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5138 GTGGGGGCTATGGCGCGGAAGAGCTACCGGAGGCGGCGGCGGAGCG 5187

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
5188 GCGAGTGGAGCGGCTATGGCAGCGCTCGCGGCGGCGGCTGCGGAGGCTA 5237

100 ...ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
5238 CGGAGGAGCGGCGGAGCGGCGGAGCTCTCGCGGCGGCTGCGAGATCATCC 5287

115 oProHishisAspHis 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5288 AGACCTCCACCAACAC 5303

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-452-872-3

seq_documentation_block:
; Sequence 3, Application US/08452872
; Patent No. 6057298
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.
; APPLICANT: Rothnagel, Joseph A.
; APPLICANT: Greenhalgh, David A.
; APPLICANT: Yuspa, Stuart H.
; TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,872
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,777
; FILING DATE:
; APPLICATION NUMBER: 07/876,289
; FILING DATE: April 30, 1992
; APPLICATION NUMBER: Unassigned (204/144)
; FILING DATE: October 29, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 204/153
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24979 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-452-872-3

alignment_scores:
  Quality: 80.50      Length: 72
  Ratio: 1.872      Gaps: 1
  Percent Similarity: 59.722      Percent Identity: 31.944

alignment_block:
  US-09-327-750D-33 x US-08-452-872-3
  Align seg 1/1 to: US-08-452-872-3 from: 1 to: 24979

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyAr 66
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
5088 CGCGGAGGCGGCGGAGCTACGCTCAGAGGCTACGCGGCGGCGGAGCG 5137

66 gTTPAsPLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluG 83
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5138 GTGGGGGCTATGGCGCGGAAGAGCTACCGGAGGCGGCGGCGGAGCG 5187

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
5188 GCGAGTGGAGCGGCTATGGCAGCGCTCGCGGCGGCGGCTGCGGAGGCTA 5237

100 ...ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
5238 CGGAGGAGCGGCGGAGCGGCGGAGCTCTCGCGGCGGCTGCGAGATCATCC 5287

115 oProHishisAspHis 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
5288 AGACCTCCACCAACAC 5303

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-03985-3

seq_documentation_block:
; Sequence 3, Application PCT/US9303985
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.
; APPLICANT: Rothnagel, Joseph A.
; APPLICANT: Greenhalgh, David A.
; APPLICANT: Yuspa, Stuart H.
; TITLE OF INVENTION: DEVELOPMENT OF A VECTOR TO TARGET GENE
; EXPRESSION TO THE EPIDERMIS OF TRANSGENIC ANIMALS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03985
; FILING DATE: 19930428
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.

```

Page 11

```

1 REGISTRATION NUMBER: 32,714
2 REFERENCE/DOCKET NUMBER: D-5478
3 TELECOMMUNICATION INFORMATION:
4 TELEPHONE: 713/651-5325
5 TELEFAX: 713/651-5246
6 TELEX: 762829
7 INFORMATION FOR SEQ ID NO: 3:
8 SEQUENCE CHARACTERISTICS:
9 LENGTH: 24979 base pairs
10 TYPE: NUCLEIC ACID
11 STRANDEDNESS: double
12 TOPOLOGY: linear
13 MOLECULE TYPE: DNA (genomic)
14 HYPOTHETICAL: NO
15 ANTI-SENSE: NO
16 PCT-UT593-03985-3

```

alignment_scores:		
Quality:	80.50	Length: .72
Ratio:	1.872	Gaps: 1
Percent Similarity:	59.722	Percent Identity: 31.944

alignment block:

US-09-327-750D-33 x PCT-US93-03985-3

Align seq 1/1 to: PCT-US93-03985-3 from: 1 to: 24979

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
||||| |||||:::|||||:::
5088 CGCGGGAGGGGGCGCAGCTACGGCTCAGGAGGCTACGGCGCGGCACGG 5137

66 gTrpAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluG 83
 5138 GTGGGGGCTATGGCGCGGAAGAAGCTACCGCGGAGCGGGGCGACGAGGC 5187

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
::: ::|||::: :: |||::: ::|||::: ::|||:::
5188 GGGAGTGGAGGGCGGTATGCCACGGCGCTCGGGCGGGTGCGGGAGCTA 5237

```

100 ...ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
      |||||
5238 CGGAGGGAGGGCAGAACGGCCCGGGATCCTCGCGGCTGCAGATCATCC 5287

```

115 OPROH1SH1SAsPH1S 120
| |||:|
5288 AGACCTCCACCAAC 5303

OM of: US-09-327-750D-33 to: N_Geneseq_1101.* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODE=frame+ p2n.spool -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09327750/runat.11032002.101154.20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -QFW=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09327750 -CGN1_1330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-33
Query length: 128
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:	Strd Orig	ZScore	Escore	Len	Documentation
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810	+	468.00	903.97	767	2.8e-42
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAI58581	+	468.00	902.84	862	3.3e-42
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611	+	468.00	902.45	898	3.4e-42
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143	+	466.50	900.72	792	4.3e-42
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	455.00	879.92	662	6.2e-41
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	455.00	879.92	698	6.2e-41
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAI60367	+	455.00	877.41	858	8.5e-41
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	444.50	858.92	692	9.1e-40
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23529	+	173.50	352.30	891	1.4e-11
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528	+	170.00	318.90	700	1.7e-09
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750	+	170.00	315.51	865	1.7e-09
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750	+	170.00	315.51	1229	1.7e-09
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	158.50	295.80	917	2.1e-08
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	135.50	255.99	532	3.5e-06
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	108.50	203.98	485	0.0028
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	106.50	201.42	421	0.0038
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	106.50	200.63	457	0.0042
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	100.50	197.49	187	0.0063
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH43880	+	96.00	180.18	451	0.0584
/SID22/gcgdata/geneseq/geneseq/NA1989.DAT:AAH90292	+	93.50	169.24	844	0.2377
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAH20086	+	91.00	141.67	8854	8.16
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82586	+	88.50	166.04	426	0.3584
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82586	+	86.00	146.63	1917	4.32
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH14000	+	86.00	145.40	2178	5.06
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH25434	+	85.00	144.15	2024	5.94
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH81776	+	85.00	137.85	3886	13.31
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58637	+	85.00	136.20	4612	16.45
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58638	+	85.00	136.14	4642	16.59
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH74810	+	85.00	135.99	4716	16.91
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH9344	+	84.00	153.43	631	1.81
/SID22/gcgdata/geneseq/geneseq/NA1997.DAT:AAH74547	+	83.50	128.60	7481	43.64
/SID22/gcgdata/geneseq/geneseq/NA1989.DAT:AAH2408	+	83.00	140.05	2061	10.04
/SID22/gcgdata/geneseq/geneseq/NA1995.DAT:AAH03508	+	82.50	150.87	607	1.26
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH3799	+	82.00	143.74	626	2.62
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH41767	+	82.00	132.48	3688	0.7552
/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAI54373	+	81.00	160.43	166	2.3355
/SID22/gcgdata/geneseq/geneseq/NA1996.DAT:AAI14335	+	81.00	149.76	502	2.89
/SID22/gcgdata/geneseq/geneseq/NA1996.DAT:AAI14335	+	81.00	147.66	624	3.79
/SID22/gcgdata/geneseq/geneseq/NA1998.DAT:AAH61594	+	81.00	145.40	789	5.06

/SID22/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ86694 - 81.00 139.01 11.48 1530
/SID22/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ86693 - 81.00 131.00 32.05 3508
/SID22/gcgdata/geneseq/geneseq/NA1992.DAT:AAQ28987 + 80.50 134.81 19.68 2137
/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ43769 + 80.50 134.51 20.45 2204
/SID22/gcgdata/geneseq/geneseq/NA1999.DAT:AAH60209 + 80.50 111.09 412.18 2497

seq_name: /SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810

seq_documentation_block:

ID_AAH75810 standard; cDNA; 767 BP.

AC_AAH75810;

DT 17-OCT-2001 (first entry)

DE Human X chromosome linked gene expression protein 14 coding sequence.

KW Human; X chromosome linked gene expression protein 14; cancer;

KW HIV infection; cytosolic; anti-HIV; chromosome X; ss.

XX Homo sapiens.

XX CN1296969-A.

XX 30-MAY-2001.

XX 23-NOV-1999; 99CN-0124078.

XX 23-NOV-1999; 99CN-0124078.

XX (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

XX Mao Y, Xie Y;

XX WPI; 2001-483897/53.

XX P-PSDB; AAG66407.

XX Polypeptide-human X chromosome linked gene expression protein 14 and

XX polynucleotide for coding said polypeptide

XX Claim 6; Page 24 (Disclosure); 3lpp; Chinese.

XX The present sequence is the coding sequence for human X chromosome linked

XX gene expression protein 14. The protein and coding sequence are useful

XX for treating diseases e.g. cancer and HIV infection.

XX Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;

alignment_scores:

Quality: 468.00 Length: 129

Ratio: 4.000 Gaps: 2

Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:

US-09-327-750D-33 x AAH75810

Align seg 1/1 to: AAH75810 from: 1 to: 767

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16

101 ATGGAGTCCAAAGAGGACGAGCGCTTAAACAATCTCATCTGCGGAAATGT 150

16 phsGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33

151 CAACACGAGAAATGATGAAAAGATGAAAGGACGACAGTGTCTAATAAG 200

33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49

201 GGGAGCCCTTGGCCCTAGCT...TTGATGTTAGTAACTACTGTGCGCT 247

50 AtgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66

248 AGAGAAACGTTAGCGGCTTCGCGCTTAGGAGCCATCCTCGAGTAG 297
 66 gTTPAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
 298 ATGGGACATATGTCATGCTTGGAGAGCCACAGGCAAGGATGAGAGAGG 347
 83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
 348 AGAATATGGAAGGATTTGGGAGGAGGTGACACAGCTGATGGAAGCTG 397
 100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 398 AGGAAAGACGATTTGAGTCTATAGTCTGCGGAGTCAGCACTGATCCCC 447
 116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
 448 TCACCATGACCATCAGGATGAGTTTGGCTTATGCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI58581

seq_documentation_block:

ID AAI58581 standard; CDNA; 862 BP.

XX AC AAI58581;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 784.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-052317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX PF WPI: 2001-442253/47.

XX DR P-PSDB; AAM39425.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries -

XX PS Claim 1; SEQ ID NO 784; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:

Quality: 468.00 Length: 129

Ratio: 4.000 Gaps: 2

Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:

US-09-327-750D-33 x AAI58581

Align seg 1/1 to: AAI58581 from: 1 to: 862

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16

226 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCGTGGAATGT 275

16 pHisGlnLysGluGluLysGluGluLysProGlnAspThrIleLysA 33

276 CAACGAGAAATGATGATAAAGATGAAAGGAGGCAAGTTGCTAATAAG 325

33 rgGluProValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49

326 GGGAGCCCTTGGCCCTACCT...TTCAATGTTAGTGAATACGTGTGCCT 372

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyra 66

373 AGAGAAACCGTAGGCGGTTCCGCGTTAGGAGCCATCTTCAGTATAG 422

66 gTTPAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83

423 ATGGACATAATGATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 472

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99

473 AGAATATGGAAGGATTTGGGAGGAGGTGAGACGCTGATGAAAGCTG 522

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116

523 AGGAAAGCAGTTGAGTCATAGTTTTCGCGGAGTCAGCACTGATCCCCC 572

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128

573 TCACCATGACCATCAGGATGAGTTTGGCTTATGCC 609

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611

seq_documentation_block:

ID AAF59611 standard; CDNA; 898 BP.

XX AC AAF59611;

XX DT 24-APR-2001 (first entry)

XX DE Human cell cycle and proliferation protein CCYPR-22 CDNA, SEQ ID NO:76.

XX KW Cell cycle and proliferation protein; CCYPR; human; agonist;

XX KW antagonist; gene therapy; detection; gene therapy;

XX KW transgenic animal disease model; immune disorder;

XX KW developmental disorder; cell signalling disorder;

XX KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;

XX KW arteriosclerosis; asthma; allergy; diabetes mellitus;

menstrual cycle disorder; bacterial infection; ss.
 XX Homo sapiens.
 PN WO200107471-A2.
 XX
 PD 01-FEB-2001.
 XX
 PF 21-JUL-2000; 2000WO-US19948.
 XX
 PR 21-JUL-1999; 99US-0145075.
 PR 08-SEP-1999; 99US-0153129.
 PR 10-NOV-1999; 99US-0164647.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman D;
 PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
 DR P-PSDB; AAB60474.
 XX
 DR WPI: 2001-112727/12.
 XX
 DR P-PSDB; AAB60474.
 XX
 PT Human cell cycle and proliferation proteins and polynucleotides are
 PT used to treat, diagnose and prevent immune, developmental and cell
 PT signaling disorders and cell proliferative disorders including cancer -
 XX
 PS Claim 5; Page 181-182; 205pp; English.
 XX
 CC Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
 CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in drug screening methods to
 CC identify compounds that modulate the activity of CCYPR. CCYPR
 CC nucleotides can be used to generate transgenic animal models of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC Diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.
 XX
 SQ Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

alignment_scores:
 Quality: 458.00 Length: 129
 Ratio: 4.000 Gaps: 2
 Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:

US-09-327-750D-33 x AAF59611 ..

Align seg 1/1 to: AAF59611 from: 1 to: 898

1 MetGluSerLysAspGlnGlyVala...LysAsnLeuAsnMetGluAsnAs 16
 254 ATGGAGTCCAAAGAGGACGCGCTTAACAATCTCATCTCGTGGAAATGT 303
 16 pHIsGlnLysLysGluGluGluLysGluGluLysProGlnAspThrIleLys 33
 304 CAACACAGGAAATGATGAAAGAGTGAAGAGGACCAAGTTCCTAATAAG 353
 33 rGluProValValalaProThrPheGluAlaGlyLysAsnCysAlaPro 49

|||||..... ||| |||..... |||..... |||
 354 GGGAGCCCTGGCCCTACCT...TTGAATGTTAGTGAATACTGTGTGCT 400
 50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
 |||||..... ||| |||..... |||..... |||
 401 AGAGAAACCGTAGCGGTTCGCGTTAGGAGCCCATCTCTGCAGTATAG 450
 66 gTTPAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
 |||||..... ||| |||..... |||..... |||
 451 ATGGGACATAATGCATAGCTTGGAGAGCCACAGCAGGAGGATGAGAGG 500
 83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
 |||||..... ||| |||..... |||..... |||
 501 AGAATATGGAAGAGGATGGGAGGAGGTGAGACAGCTGATGGAAAGCTG 550
 100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 |||||..... ||| |||..... |||..... |||
 551 AGGAAACAGCTTGAGTCATAGTCTGCGGCGAGTCAGCACTGATCCCC 600
 116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
 |||||..... ||| |||..... |||..... |||
 601 TCACCATGACCATCAGATGAGTGTTCCTTATGCC 637

seq name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard; cDNA; 792 BP.

XX AC AAH45143;

XX AC

XX DT 07-SEP-2001 (first entry)

XX DE Human brain expressed X-linked protein, hbex, coding sequence.

XX KW Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;

XX KW hbex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;

XX KW X-chromosome-binding mental retardation; lissencephalous disease; ss.

XX OS Homo sapiens.

XX PN WO200140286-A1.

XX PD 07-JUN-2001.

XX PF 27-NOV-2000; 2000WO-CN00502.

XX PR 30-NOV-1999; 99CN-0124179.

XX PA (BIOR-) BIOROAD GENE DEV LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX DR WPI: 2001-397944/42.

XX DR P-PSDB; AAB99224.

XX PT Isolated human brain-expressed X-linked polypeptide used to diagnose
 PT and treat of dysembryoplasia, hereditary diseases, cancer, tumour,
 PT deafness and X-chromosome-binding mental retardation

XX PS Claim 5; Page 22; 30pp; Chinese.

XX CC The present sequence is the coding sequence for a human brain-expressed
 CC X-linked protein (hbex). hbex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, X-chromosome-binding mental retardation and
 CC lissencephalous disease. hbex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hbex coding sequence can be used as primers or probes, or in producing
 CC gene chips or microarrays.

XX SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 3877.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
XX (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX P-PSDB; AAG03873.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 3877; 71pp + CD-ROM; English.
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or poly(A) RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dr primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX SQ Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other:

alignment_scores:
Quality: 455.00 Length: 129
Ratio: 4.027 Gaps: 2
Percent Similarity: 87.597 Percent Identity: 66.667

alignment_block:
US-09-327-750D-33 x AAC03879 ...
Align seg 1/1 to: AAC03879 from: 1 to: 698

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
243 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCGTGGAAATGT 292
16 PHisGlnLysGluGluGlnGlyLysProGlnAspThrLeuLysA 33
293 CAACAGAAATGATGAAATGATGAAATGAAAGGAGCAAGTGGCTATTAAG 342
33 rGluProValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
343 GGGAGCCCTTGGCCCTACCT...TTGRATGYTRGTGAATACACTGTGCT 389
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
390 AGAGGAATCGTAGGCGGTTCGCCGTTAGGCAGCCCATCTCCAGCATATAG 439

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
440 ATGGGATATGATGATAGGCTGGAGAACACACAGGCAAGGATGAGAGAAG 489
83 LuAsnValClnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
490 ACAATATGAAAGGATTGGGAGGAGGAGTGGAGACAGCTGATGAGAAAGCTG 539
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
540 AGGGAAAGACAGTGTAGTCATAGTCTCGGGCAGTCAGCACTGACCCCCC 589
116 OHishHisAspHisHisAspGluPheCysLeuMetPro 128
590 TCACCATGACCATCATGATGAGTTTTCNNNATGCCCC 626
seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAI60367
seq_documentation_block:
ID AAI60367 standard; cDNA; 858 BP.
XX AC AAI60367;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 4356.
XX KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-052317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;
XX WPI; 2001-442253/47.
XX DR P-PSDB; AAM41211.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 4356; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 858 BP; 236 A; 175 C; 241 G; 206 T; 0 other;

alignment_scores:
 Quality: 455.00 Length: 130
 Ratio: 3.889 Gaps: 3
 Percent Similarity: 90.000 Percent Identity: 67.692

alignment_block:

US-09-327-750D-33 x AAC160367

Align seg 1/1 to: AAC160367 from: 1 to: 858

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
 215 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCGTGGAAAATGT 264
 16 PHisGlnLysLysGluGlnLysGluGlnLysProGlnAspThrIleLysA 33
 265 CAACGAGAAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTATAAAG 314
 33 tGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
 315 GGGAGCCCTGGCCCTACCT...TTGAATTTAGTGAATCTGTGCTGCT 361
 50 ArgGlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyraR 66
 362 AGAGAAACCCGTAGCGGTTCCGCGTTAGGACGCCCATCTCGCAGTATAG 411
 66 gTIPAspLeuMethisArgValGlyGluProGlnGlyArgMetArgGluG 83
 412 ATGGACATAATGTCATAGGCTTGAGAGCCACAGGCAAGGATGAGAGAGG 461
 83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
 462 AGAATATGGAAGGATTTGGGAGAGGTGAGACAGCTGATGGAAGCTG 511
 100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 512 AGGGAAGAGCAGTTGAGTTCATAGTCTCGGGCAGTCAGCACTGATCCCCC 561
 116 oHisHisAspHisAspGluPheCys.LeuMetPro 128
 562 TCACCATGACCATCAGATGAGTTTGCCTTTATGCC 599

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2000.DAT:AAC10889

seq_documentation_block:

ID AAC10889 standard; cDNA; 692 BP.

XX AAC10889;

AC AAC10889;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 14964.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

OS Homo sapiens.

XX EP1033401-A2.

XX

PD 06-SEP-2000.
 XX 21-FEB-2000; 2000EP-0200610.
 PF 26-FEB-1999; 99US-0122487.
 PR (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 FI WPI; 2000-500381/45.
 DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT Claim 1: SEQ ID 14964; 71pp + CD-ROM; English.
 PS The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX SQ Sequence 692 BP; 184 A; 151 C; 199 G; 146 T; 12 other;

alignment_scores:
 Quality: 444.50 Length: 125
 Ratio: 4.041 Gaps: 1
 Percent Similarity: 88.000 Percent Identity: 64.800

alignment_block:

US-09-327-750D-33 x AAC10889

Align seg 1/1 to: AAC10889 from: 1 to: 692

4 LysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHisGlnLysLy 20
 249 AGAGAACGAGCGTTAAACAATCTCATCGTGGAAAATGTCAACCGAGAAA 298
 20 sGluGluLysGluGluLysProGlnAspThrIleLysArgGluProValv 37
 299 TGATGAAAAGATGAAAAGGAGCAAGTGTCTATAAAGGGGAGCCCTGG 348
 37 alAlaProThrPheGluAlaGlyLysAsnCysAlaProArgGlyGlyArg 53
 349 CCTACCT...TTGATGTTGTGAATCTGTGCTAGAGGAAATCGT 395
 54 ArgArgPheArgValArgGlnProIleSerHisTyraArgTIPAspLeuMe 70
 396 AGGCGGTTCCGCGTTAGGACGCCCATCTCGCAGTATAGTGGATATGAT 445
 70 tHisArgValGlyGluProGlnGlyArgMetArgGluAsnValGlnA 87
 446 GCATAGGCTTGGAGAACCAACAGGCAAGGATGANAAGAGAATATGGAAA 495
 87 rgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGln 103
 496 GGATTGGGAGAGGTGAGACAGCTGATGGAAGCTGAGGGAAGACAG 545
 104 LeuSerHisSerLeuArgAlaValSerThrAspProHisAspHis 120
 546 TTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCCATCACCATGACCA 595

120 shisaspGluPheCysLeuMetPro 128
 |||||
 596 TCATGATGAGTTTGCNNWATGCC 620

seq_name: /SIDS2/gcgdata/geneseq/NA2001.DAT:AAF23529

seq_documentation_block:

ID AAF23529 standard; DNA; 891 BP.

XX AAF23529;

XX 22-MAR-2001 (first entry)

XX Human NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
 KW neurogenetic disease; NF-kappaB; ds.

XX Homo sapiens.

XX WO200075278-A2.

XX 14-DEC-2000.

XX 07-JUN-2000; 2000WO-US15621.

XX 07-JUN-1999; 99US-0327750.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Sato T;

XX WPI; 2001-061707/07.

XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
 PT the gene encoding NADE, useful for modulating the activity of p75NTR
 and for detecting neurodegenerative diseases .

XX Disclosure; Fig 1; 134pp; English.

XX The present invention relates to a purified polypeptide capable of
 CC binding neurotrophin receptor (p75-NTR). The invention is useful for
 CC binding and modulating the activity of p75NTR. The peptide mediates
 CC NGF-induced apoptosis, which plays an important role in neurogenetic
 CC diseases. The peptide of the invention and p75NTR are useful for
 CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
 CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
 CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.

XX Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

alignment_scores:

Quality: 187.50 Length: 129
 Ratio: 2.232 Gaps: 5
 Percent Similarity: 65.116 Percent Identity: 36.434

alignment_block:

US-09-327-750d-33 x AAF23529 ..

Align seq 1/1 to: AAF23529 from: 1 to: 891

2 GluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHisG1 18

279 AAGAAACACACACAGAGAAAAAATCTCATCATGCGCAATATTCACCA 328

18 nLysLysGluGluLysProGlnAspThrIleLysArgGluP 35

329 GGAACACGAGATGGAGCGCTATGCAGATGGAGAGAGACCGCC 378

35 roValAlaProThrPheGluAlaGlyLysAsnCysAlaProArgGly 51

379 CTTTG.....GGAGAGGTGAAGCCACCAGCCT...GCA 410
 52 GlyArgArgPheArgValArgGlnProIleSerHisTyrArgTrpAs 68
 |||||
 411 GGAATATCGACGGGACAGGCTCGCGACTTCCCTTAATTTTCATGGGC 460
 68 pLeuMetHisArg...ValGlyGluProGlnGlyArgMetArgGluGluA 84
 :|||
 461 CATACCAATAGCAGATCAATGATGGATGGTGA...GATGGAGATG 507
 84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
 :|||
 508 ATATGCAATATTCATGAGGAGATGAGAGAAATCAGAGAAACTTAGG 557
 101 GluArgGluLeuSerHisSerLeuArgAlaValSerThrAsp...ProPr 116
 |||||
 558 GAGTCGAGTTGAGGAATGTCGCTATCCTTATGGGAGGCTCTCTAA 607
 116 ohHisAspHisHisAspGluPheCysLeuMetPro 128
 |||||
 608 TCACCATGACCATCATGATGAATTTTCCTTATGCCT 644

seq_name: /SIDS2/gcgdata/geneseq/NA2001.DAT:AAF23528

seq_documentation_block:

ID AAF23528 standard; DNA; 700 BP.

XX AAF23528;

XX 22-MAR-2001 (first entry)

XX Mouse NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
 KW neurogenetic disease; NF-kappaB; ds.

XX Mus sp.

XX WO200075278-A2.

XX 14-DEC-2000.

XX 07-JUN-2000; 2000WO-US15621.

XX 07-JUN-1999; 99US-0327750.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Sato T;

XX WPI; 2001-061707/07.

XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
 PT the gene encoding NADE, useful for modulating the activity of p75NTR
 and for detecting neurodegenerative diseases .

XX Claim 12; Fig 1; 134pp; English.

XX The present invention relates to a purified polypeptide capable of
 CC binding neurotrophin receptor (p75-NTR). The invention is useful for
 CC binding and modulating the activity of p75NTR. The peptide mediates
 CC NGF-induced apoptosis, which plays an important role in neurogenetic
 CC diseases. The peptide of the invention and p75NTR are useful for
 CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
 CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
 CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.

XX Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

alignment_scores:

Quality: 173.00 Length: 139
 Ratio: 2.190 Gaps: 5

Percent Similarity: 56.835 Percent Identity: 32.374

alignment_block:

US-09-327-750D-33 x AAF23528

Align seg 1/1 to: AAF23528 from: 1 to: 700

```
9 LysAsnLeuAsnMetGluAsnAspHisGlnLysLysGluGluLysGluGlu 25
   ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
165 CAAATCTCATCTGCGGCAATGTCCACGAGAAAGAGAGCTGGAGCA 214
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 uLysProGlnAspThrIleLysArgGluProValValAlaProThrPheG 42
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 GCCCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
42 LuAlaGlyLysAsnGlnAlaProArgGlyGly 52
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 GAGGAGGTGAGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 299
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
53 ....., ArgArgArgPheArgValArgGlnProIleLeu 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 AACCATACCAACACACACACACACACACACACACACACACACACACAC 349
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 rHisTyrArgTrpAspLeuMetHisArgValGlyGluProGlnGlyArgM 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 TAACTTCGATGGGCCATT.....CCCAACAGGCAGA 381
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
80 eArgGlu.....GluAsnValGlnArgPheGlyGlu 90
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 TGAATGACGGGTGGTGGGAGATGATGATGATGATGATGATGATGATGAT 431
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 AspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSerHisLe 107
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
432 GAGATGAGAGATCCGAGAGAAAGCTTAGGGAGCTACAGCTGAGAAATG 481
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
107 rLeuArgAlaValSerThrAsp...ProHisHisAspHisHisAspG 123
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
482 TCACGCATCTTTATGGGGAGGTGTCTAACCCAGCATCACCATTGATG 531
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 LuPheCysLeuMetPro 128
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 AATTCTGCCTTATGCT 548
```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID: AAH03517 standard; cDNA; 865 BP.

XX AC AAH03517;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:352.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
of the abnormality of the proteins encoded by the
full-length cDNAs -

XX PS Claim 1; SEQ ID 352; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dT primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH2446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.

XX SQ Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:

Quality: 170.00 Length: 130

Ratio: 1.954 Gaps: 5

Percent Similarity: 66.923 Percent Identity: 33.077

alignment_block:

US-09-327-750D-33 x AAH03517

Align seg 1/1 to: AAH03517 from: 1 to: 865

```
1 MetGluSerLysAspGln...GlyAlaLysAsnLeuAsnMetGluAsnAs 16
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 ATGGAGTCCAAAGAGAACTAGCGGCAACAATCTCAACGGGAAATGC 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16 phisGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 CCAACAAGAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 290
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 AATCCCGCCATTGGGA.....GGGGGTGAAGCCAGAGAGCCT 328
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
329 GGAGGAATATACGCGGGGGGAGTAGGCGACTTGTCCCTAATTTCG 378
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 gTrpAspLeuMetHisArg...ValGlyGluProGlnGlyArgMetArg 82
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 ATGGGCCATACCTATAGCATATTGAGCACAATGAAGCGAGA..... 421
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82 LuGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLys 98
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
422 ..GATGATGAGAAAGGTTTGTAGGCAGATGATGGAATCAAGAGAAAG 469
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 LeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

alignment_block:
US-09-327-750D-33 x AAF21748
Align seg 1/1 to: AAF21748 from: 1 to: 917
18 GlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysArgG1 34
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
306 CAGGAAACGAGAGATGGAGCAGCCTATCGAAGTGGAGAGAACGCG 355
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 uProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
356 CCCTTGG.....GGAGGAGTGAAGGCCACCGCT...G 387
51 LyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTTP 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
388 CAGGAAATCGAGGGGACAGCGCTCGCGACTTGCCTTAATTTTCGATGG 437
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
68 AspLeuMetHisArg...ValGlyGluProGlnGlyArgMetArgGlu 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
438 GCCATACCAATAGGCAGATCAATGATGGATGGTGA...GATGGAGA 484
83 uAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
485 TCATATGGAATATTCATGGAGGAGATGAGAAATCAGAGAAACTTA 534
100 rgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp...Pro 115
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
535 GGGAGCTGCAGTTCAGGAATTCCTGCGTATCCTTATGGGGAGCTCTCT 584
116 ProHisHisAspHisHisAspGluPheCysLeuMetPro 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
585 AATCACCATGACATCATGATGATTTTGCTTATGCGCT 623

seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA2000.DAT:AAC01005
seq_documentation_block:
ID AAC01005 standard; cDNA; 532 BP.
XX
XX AAC01005;
XX
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 1003.
XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX
XX EPI033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
XX
XX P-PSDB: AAG00999.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the

```

```

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
XX Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;

```

```

alignment_scores:
Quality: 135.50 Length: 120
Ratio: 1.831 Gaps: 5
Percent Similarity: 61.667 Percent Identity: 32.500
alignment_block:
US-09-327-750D-33 x AAC01005
Align seg 1/1 to: AAC01005 from: 1 to: 532
4 LysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHisGlnLysLy 20
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
192 AAACAACCCAGAAAAAATCTCATCATGCAATATTCACCAAGAAAA 241
20 sGluGluLysGluGluLysProGlnAspThrIleLysArgGluProVal 37
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
242 CGAAGAGATGGAGCAGCCTATCCAGATGGAGAGAACGCCCTTTG. 290
37 alAlaProThrPheGluAlaGlyLysAsnCysAlaProArgGlyGlyArg 53
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 .....GGAGGAGTGAAGGCCACCGCT...GCAGGAAT 323
54 ArgArgPheArgValArgGlnProIleSerHisTyrArgTTPAspLeuMe 70
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
324 CGACGGGAGAGGCTCGCGGAYTTGCCCTTAATTTTCGATGGCCATACC 373
70 tHisArg...ValGlyGluProGlnGlyArgMetArgGluAsnValG 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
374 CAATAGGCAGATCAATGATGGATGGGT...GRAGATGGAGATGATATGG 420
86 lnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArgGluArg 102
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 AAATATTCATGGAGAGATGAGAAATCAGAAACAACTTAGGAGCTG 470
103 GlnLeuSerHisSerLeuArg...AlaValSerThrAspProHisHl 118
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 CAKTTGAGGAATTCCTGCTATCTTATGGGGAKTCTCTTAATACCA 520
118 sAspHisHis 121
||||| |||||
521 TGACCATCAT 530

```

	+	91.00	115.68	362.85	17992	I ASQ07125 Mycobacterium tub
gb_gb:ASQ07125	-	91.00	111.49	620.87	31175	I U00034 Mycobacterium tuber
gb_gb_ba:MSU0024	-	91.00	111.49	620.89	31176	I A0000009 Mycobacterium tub
gb_gb_ba:MNGY2	-	91.00	108.95	860.19	43523	I ZB3018 Mycobacterium tuber
gb_gb_ba:MTSGY349	-	91.00	108.95	860.19	43523	I ZB3018 Mycobacterium tuber
gb_gb_ba:MTCY349	-	91.00	108.95	860.19	43523	I ZB3018 Mycobacterium tuber
gb_htg:AP001332	+	91.00	100.23	2.6e+03	136678	I AP001332 Homo sapiens chr

seq_documentation_block:
LOCUS AF097437 2269 bp DNA 05-AUG-1999
DEFINITION Mus musculus Bex1 protein (Bex1) gene, complete cds.
ACCESSION AF097437
VERSION AF097437.1 GI:5702155
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 2269)
Brown,A.L. and Kay,G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619(1999)
JOURNAL

is a gene with increased expression in pancreatic carcinoma. This gene is a member of a novel gene family on the mouse X chromosome. Hum. Mol. Genet. 8 (4), 611-619 (1999)

99172070
PUBMED 10072429
2 (bases 1 to 2269)
Brown, A.L. and Kay, G.F.
Direct Submission
Submitted (07-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4030, Australia

FEATURES
source 1
1..2269

Strd	Orig	zScore	EScore	Len	Documentation
+	646.00	986.65	1.1e-46	3269	! AF097437 Mus musculus Bx1 prc
+	640.00	985.02	1.4e-46	835	! AF097438 Mus musculus brain exp
+	640.00	984.99	1.4e-46	838	! AF051347 Mus musculus REX-3 mRN
+	601.00	926.17	2.6e-43	785	! AF097439 Mus musculus brain exp
+	582.00	886.16	4.1e-41	2901	! BC003254 Mus musculus, Similar
+	471.50	727.38	3.1e-32	642	! AF183416 Homo sapiens ovarian g
+	471.50	725.44	3.9e-32	828	! AF220189 Homo sapiens uncharac
-	471.50	725.44	3.9e-32	35714	! AL008708 Homo sapiens DNA sequen
+	468.00	720.87	1.6e-32	744	! AF251053 Homo sapiens X-linked
+	468.00	719.43	8.5e-32	898	! AF078272 Homo sapiens 76 from Paten
+	468.00	692.37	2.7e-30	31321	! Z70233 Human DNA sequence fr
-	468.00	690.40	3.5e-30	40584	! AL133348 Homo sapiens DNA sequen
+	468.00	690.32	3.6e-30	41029	! AL022169 Homo sapiens chromos
+	466.00	718.09	1.0e-31	791	! AF377783 Homo sapiens brain-exp
-	311.00	481.91	1.4e-18	504	! G24641 Human STS WI-11354, sequ
+	195.50	269.60	9.7e-07	45678	! Z70719 Human DNA sequence fr
+	187.50	287.27	1.0e-07	891	! M31818 Human unknown protein fr
+	187.50	287.27	1.0e-07	891	! AF187064 Homo sapiens, p75NTR-as
+	186.50	286.62	1.1e-07	793	! BC003190 Homo sapiens, p75NTR-as
+	174.00	268.30	1.1e-06	700	! AF187066 Homo sapiens, p75NTR-as
-	174.00	266.79	1.4e-06	854	! AF097440 Mus musculus brain exp
-	170.00	265.06	1.7e-06	477	! G35294 human STS SHGC-37409, se
+	170.00	257.85	4.4e-06	1229	! AK000959 Homo sapiens cDNA FLJ
+	170.00	257.06	4.8e-06	1364	! AX100231 Sequence 32 from Paten
+	170.00	227.10	0.0002	69648	! AL035494 Human DNA sequence f
+	148.00	230.52	0.0001	519	! AF187065 Rattus norvegicus p75N
-	142.50	223.64	0.0004	421	! G72708 MARC 4953-4954:99139031
+	102.00	127.34	81.34	36028	! AL442120 Streptomyces coelic
-	100.50	159.86	1.25	372	! G23964 human STS WI-15922, sequ
+	97.50	136.41	25.41	4408	! AB040971 Homo sapiens mRNA for
+	94.50	103.23	1.8e+03	187075	! AC021617 Homo sapiens clone
+	93.50	142.84	11.15	844	! A25464 B.pertussis STX pillin ge
+	93.50	142.83	11.15	845	! Y00556 Bordetella pertussis fim
+	93.50	139.46	17.18	1315	! X74118 B.bronchiseptica fimx g
+	93.00	139.46	17.18	1315	! AF3843 Sequence 1 from Patent
+	93.00	130.97	51.07	3624	! AR029509 Sequence 6 from patent
+	93.00	130.97	51.07	3624	! AR098462 Sequence 6 from patent
+	93.00	130.97	51.07	3624	! I41422 Sequence 6 from patent
-	92.00	133.28	37.94	2184	! L20442 Rhesus monkey p53 mRNA,
-	92.00	122.97	142.38	8455	! U95372 Halofret volcanici blas

```

/organism="Mus musculus"
/strain="129SVJ"
/db_xref="taxon:10090"
/chromosome="X"
/map="near Plp"
/tissue_type="liver"
/gene="Bex1"
/product="Bex1 protein"
409..1873
/gene="Bex1"
/note="expressed in brain; X-linked"
1268..1654
/gene="Bex1"
/codon_start=1
/product="Bex1 protein"
/protein_id="AA047168.1"
/db_xref="GI:5702156"
/translation="MESKQDGVKNI.NMENDHKKKEEKPODTIRREPAVALTSG
KNCAPGRRRRVQPTAHYNWDLQMVGPEQPGQREENVQRFGGDVRLMEKLRER
QLSHSLRAYSTDPHPHDDHDEFCLMP"
563 a 554 c 705 g 447 t
BASE COUNT
100

```

alignment_scores:		
Quality:	646.00	Length: 128
Ratio:	5.168	Gaps: 0
Percent Similarity:	97.656	Percent Identity: 93.969

```
alignment_block:
US-09-327-750D-33 x AF097437
```

Align seq 1/1 to: AF097437 from: 1 to: 2269

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHI 17
 2626 ATGGAGTCCAAAGATCAAGCGGTGAATAATCTCAACATGGAGATGACCA 1317
 17 sGlnLysLysGluLysGluLysGluLysProGlnAspThrLysArgC 34


```

/note="retinoic acid reduced gene expression"
/codon_start=1
/product="REX-3"
/protein_id="AAC61929.1"
/db_xref="GI:3510643"
/translation="MENDHOKKEEKEKPODTRIRPAVALISEAGKNCAPRGRRRF
RVQPIAHYRDLQVORVGEPOGRMREENVORFGDVRQLMEKLRQLSHSLRAVSTD
PPHHHDEFLCP"
BASE COUNT      238 a 181 c 251 g 168 t
ORIGIN

alignment_scores:
  Quality: 640.00      Length: 128
  Ratio: 5.161         Gaps: 0
Percent Similarity: 96.875 Percent Identity: 92.188.

alignment_block:
US-09-327-750D-33 x AF051347

Align seg 1/1 to: AF051347 from: 1 to: 838

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH 17
|||||
193 ATGGAGTCCCAAGATCAAGCGGTGAAATAATCTCAACATGAGATGACCA 242

17 sGlnLysLysGluLysGluLysProGlnAspThrIleLysArg 34
|||||
243 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 292

34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
293 AGCCAGCTGTGGCCCTGATCTCCAGGCTGGCAAAACCTGCGCCCTAGA 342

51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
343 GGAGGTCGACGGGTTCGGGTTCGGCAGCCATCGCTACATATAGATG 392

67 pasLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
393 GGACCTGATGAGAGGGTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 442

84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||
443 ACCTACAGAGTTTGGGGTGTGTGAGACAGCTCATGGAGAAGCTGAGG 492

101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
493 GAAGGAGCTGAGCCAGCTCGGGCGGTAGCACTGACCTGACCCGCTCA 542

117 shisAspHisAspGluPheCysLeuMetPro 128
|||||
543 TCATGACCACCATGATGATGTTTGGCTCATGCC 576

seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS      AF097439      785 bp      mRNA      13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 2 (Bex2) mRNA,
complete cds.
ACCESSION  AF097439
VERSION    AF097439.1 GI:4580591
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM  Mus musculus
REFERENCE  1 (bases 1 to 785)
AUTHORS   Brown,A.L. and Kay,G.F.
TITLE     Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL   Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE   99172070

Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 785)
AUTHORS   Brown,A.L. and Kay,G.F.
TITLE     Direct Submission
JOURNAL   Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
Location/Qualifiers
FEATURES
Source
1..785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:W48832"
/chromosome="X"
/map="near plp"
/dev_stage="embryo; 15.5 dpc"
1..785
/gene="Bex2"
/locus_tag="Bex2"
/codon_start=1
/product="brain expressed x-linked protein 2"
/protein_id="AAD24430.1"
/db_xref="GI:4580592"
/translation="MESKYEQGVKLNNDHOKKEEKEKPODASKRDPITVALPFEA
QYIVPRGRRRRFRQPIVHYRDLMEHVRVGEPOGRMREENVORFGDVRQLMEKLR
QLYSHSLRAVSTDPPHHHDEFLCP"
BASE COUNT      212 a 163 c 235 g 175 t
ORIGIN

alignment_scores:
  Quality: 601.50      Length: 129
  Ratio: 4.971         Gaps: 1
Percent Similarity: 93.798 Percent Identity: 86.822

alignment_block:
US-09-327-750D-33 x AF097439

Align seg 1/1 to: AF097439 from: 1 to: 785

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16
|||||
139 ATGGAGTCCCAAGATGGAACAGCGGTGAAATAATCTCAACATGGAATGA 188

16 pHisGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
|||||
189 CCATCAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 238

33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||
239 GGGATCCGATTTGGCCCTGCTTCGAGGCTGAGAGCTACTACGTGCT 288

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||
289 AGAGGAGGTTCGAGCGGTTCGGGTTCGGCAGCCCATCGTCACACAG 338

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83
|||||
339 ATGGGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 388

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||
389 AGAACGTACAGAGTTTGGGATGATGATGATGATGATGATGATGATG 438

100 ArgGluArgGlnLeuSerHisSerHisLeuArgAlaValSerThrAspPro 116
|||||
439 AGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 488

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
489 TCATCATGACCACCATGATGATGATGATGATGATGATGATGATGATG 525
```

seq_name: gb_ro:BC003254

seq_documentation_block: 2901 bp mRNA ROD 20-FEB-2001
 LOCUS BC003254 Mus musculus, similar to dentatorubral pallidoluysian atrophy,
 clone MGC:5758, mRNA, complete cds.
 DEFINITION BC003254
 ACCESSION BC003254
 VERSION BC003254.1 GI:13096906
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2901)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalona@bcm.tmc.edu.
 Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 9 Row: 0 Column: 6.

FEATURES

source
 1..2901
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:5758 IMAGE:3500522"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma, 5 month old virgin mouse."
 /clone_lib="NCI_CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 755..2359
 /codon_start=1
 /product="Similar to dentatorubral pallidoluysian atrophy"
 /protein_id="AAH03254.1"
 /db_xref="GI:13096907"
 /translation="MVQFLLTLPGLPKPIHVHTPLHPRAKPHPTPSHQSLQSP
 PQLPFLSSPPAAPTTPITATQIKQEAEEYEPSPVPAPSPSPPKVVDVPS
 HASQARNKHLDRGNSCARSDLYFVLEGLAKKRLADLVEKVRREARAREKE
 REREREKEREREKERSVLAQGRAPVECPSLGPPVHPPEPGSAVATVPP
 YLGPTPALTLSEVAPHPVMSPCNRNHPVPLGAVDPCLLGYNPALYSSDPAARE
 RERERERDRDLKPGFEVAPSELPSELPVLPGLDPPFRHGLALQPGPLHPFP
 FHPSPGLERELALAGPALRPMYSALRAERQHAERVAALNDPLRPLNLT
 PPHHQSHHSHLHLHQDDAIHAASAVHPLIDPLASGSHLTRIPYAGTLPLNPLP
 PLRNEVLRHQLFAYPYDPLASISAPMSAAHQIQAMHQAQSLQRLALEQQQWLHAH
 HPLHSVPLPAQEDYVSHLKESDKPL"

CDS

BASE COUNT 623 a 958 c 791 g 529 t
 ORIGIN

alignment_scores:

Quality: 582.00 Length: 118
 Ratio: 5.061 Gaps: 0
 Percent Similarity: 97.458 Percent Identity: 92.373

alignment_block:

US-09-327-750D-33 x BC003254

Align seg 1/1 to: BC003254 from: 1 to: 2901

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
 |||||
 223 ATGGAGTCCAAAGATCAAGGCGTGAAATCTCAACATGGAGATGACCA 272

17 sGlnLysLysGluGluLysProGlnLysProGlnAspThrIleLysArg 34
 |||||
 273 TCGAATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 322

34 luProValValalaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
 |||||
 323 AGCCAGCTGTGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 372

51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
 |||||
 373 GGAGTTCGACGGCGGTTCGGGTCGCGAGCCCATCGCTCACTATAGTG 422

67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluAla 84
 |||||
 423 GCACCTGATGCAGAGGGTTGGGAGAGCCCGAGGAGGAGGAGGAGGAGA 472

84 snValGluArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
 |||||
 473 AGCTACAGAGGTTGGGGGTGTGTGAGACAGCTCATGGAGAGCTGAGG 522

101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProH1 117
 |||||
 523 GAAGGACAGCTGAGCCACAGCTCGGGCGGTAGCAGCTGACCCGCGCTCA 572

117 shis 118
 |||||

573 TCAT 576

seq_name: gb_pr:AF183416

seq_documentation_block: 642 bp mRNA PRI 02-SEP-2000

LOCUS AF183416 Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
 DEFINITION A novel gene expressed in human adrenal gland
 ACCESSION AF183416
 VERSION AF183416.1 GI:9963770
 KEYWORDS human.
 SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 642)
 AUTHORS Peng, Y., Qian, B., Tu, Y., Xu, S., Han, Z., Fu, G. and Chen, Z.

TITLE A novel gene expressed in human adrenal gland
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 642)
 AUTHORS Peng, Y., Gu, W., Huang, C., Xu, S., Han, Z., Fu, G. and Chen, Z.

TITLE Direct Submission
 JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at
 Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
 Shanghai 201203, P. R. China

FEATURES
 source
 1..642
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="adrenal gland"
 3..380
 /note="HGR74-h protein"
 /codon_start=1
 /evidence="not_experimental"
 /product="ovarian granulosa cell 13.0 kDa protein hGR74
 homolog"
 /protein_id="AA09685.1"
 /db_xref="GI:9963771"

/translation="MESKEKRAVNSLSMENAOENKEEQVANKGRLPLALPLDAGEYC
VPRGNRRFRVROPILQYRWDMHRLGEPQARMRENNRIGEEVRLMEKLEKREKOLS
HSLRAVSTDPHPHHDHDFCLMP"

BASE COUNT 204 a 118 c 157 g 163 t
ORIGIN

alignment_scores:
Quality: 471.50 Length: 129
Ratio: 4.100 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 70.543

alignment_block:

US-09-327-750D-33 x AF183416 ..

Align seg 1/1 to: AF183416 from: 1 to: 642

```

1 MetGluSerLysAspGlnGlyAlaLysAsn...LeuAsnMetGluAsnAs 16
|||||
3 ATGGAGTCCAAAGAGAACGACGAGTAAACAGTCTCAGCATGGAATAATGC 52
|||||
16 phisGlnLysLysGluLysGluLysProGlnAspThrIleLysA 33
|||||
53 CAACCAAGAAATGAAGAAAGAG.....CAAGTTGCTAATAAAG 93
|||||
33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||
94 GGGAGCCCTTGGCCCTCCCT...TTGGATGCTGGTGAATCTGTGTCCT 140
|||||
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||
141 AGAGAAATCTAGCGGTTCCCGGTTAGGCACCCCATCTCGCAGTATAG 190
|||||
66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83
|||||
191 ATGGGATATGATGATAGGCTGGAGAACACACAGGCAAGGATGAGAGAAG 240
|||||
83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||
241 AGAATATGGAAGGATGGGAGGAGGTGAGACAGCTGATGGAAGAAGCTG 290
|||||
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||
291 AGGGAAGAGCAGTGTAGTCTAGTCTCGGGCAGTCAAGCACTGACCCCC 340
|||||
116 ohHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
341 TCACCATGACCATCATGATGAGTTTTCCTTATGCC 377

```

seq_name: gb_pr:AF220189

seq_documentation_block:

LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.

ACCESSION AF220189

VERSION AF220189.1 GI:7689028

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1. (bases 1 to 828)

TITLE Xiao.H., Song.H., Gao.G., Ren.S., Chen.Z. and Han.Z.

JOURNAL A novel gene expressed in human hypothalamus

REFERENCE 2 (bases 1 to 828)

AUTHORS Xiao.H., Song.H., Gao.G., Ren.S., Chen.Z. and Han.Z.

TITLE Direct Submission

JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China

FEATURES

Location/Qualifiers

source

1. .828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67654.1"
/db_xref="GI:7689028"
/translation="MESKEKRAVNSLSMENAOENKEEQVANKGRLPLALPLDAGEYC
VPRGNRRFRVROPILQYRWDMHRLGEPQARMRENNRIGEEVRLMEKLEKREKOLS
HSLRAVSTDPHPHHDHDFCLMP"

CDS

BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN

alignment_scores:

Quality: 471.50 Length: 129
Ratio: 4.100 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 70.543

alignment_block:

US-09-327-750D-33 x AF220189 ..

Align seg 1/1 to: AF220189 from: 1 to: 828

```

1 MetGluSerLysAspGlnGlyAlaLysAsn...LeuAsnMetGluAsnAs 16
|||||
167 ATGGAGTCCAAAGAGAACGACGAGTAAACAGTCTCAGCATGGAATAATGC 216
|||||
16 phisGlnLysLysGluLysGluLysProGlnAspThrIleLysA 33
|||||
217 CAACCAAGAAATGAAGAAAGAG.....CAAGTTGCTAATAAAG 257
|||||
33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||
258 GGGAGCCCTTGGCCCTCCCT...TTGGATGCTGGTGAATCTGTGTCCT 304
|||||
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||
305 AGAGAAATCTAGCGGTTCCCGGTTAGGCACCCCATCTCGCAGTATAG 354
|||||
66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83
|||||
355 ATGGGATATGATGATAGGCTGGAGAACACACAGGCAAGGATGAGAGAAG 404
|||||
83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||
405 AGAATATGGAAGGATGGGAGGAGGTGAGACAGCTGATGGAAGAAGCTG 454
|||||
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||
455 AGGGAAGAGCAGTGTAGTCTAGTCTCGGGCAGTCAAGCACTGACCCCC 504
|||||
116 ohHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
505 TCACCATGACCATCATGATGAGTTTTCCTTATGCC 541

```

seq_name: gb_pr:HS198P4

seq_documentation_block:

LOCUS HS198P4 35714 bp DNA PRI 16-DEC-2000
DEFINITION Human DNA sequence from clone RP1-198P4 on chromosome Xq22 Contains
a CpG island, complete sequence.

ACCESSION AL008708

VERSION AL008708.4 GI:4826445

KEYWORDS HTG; CpG Island.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 35714)

AUTHORS Corby,N.

TITLE
JOURNAL

Direct Submission
Submitted (21-JAN-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk

COMMENT

On May 13, 1999 this sequence version replaced g1:4592117.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/ChrX>

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RPI-198P4 is
from the library RPI-1 constructed at the Roswell Park Cancer

Institute by the group of Pieter de Jong. For further details see
<http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone

RPI-198P4 it may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

The true left end of clone LLOXNC01-221F2 is at 35615 in this
sequence. The true right end of clone LLOXNC01-73E8 is at 100 in
this sequence.

FEATURES

source

Location/Qualifiers
1..35714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22"
/clone="RPI-198P4"
/clone_lib="RPI-1"

repeat_region
1..258
/note="L1PA3 repeat: matches 5665..5925 of consensus"
256..537
/note="L1PA3 repeat: matches 5391..5672 of consensus"
1491..1749
/note="AluSc repeat: matches 39..297 of consensus"
2182..3810
/note="L1PB3 repeat: matches 4533..6150 of consensus"
5169..6687
/note="L1MA2 repeat: matches 4798..6308 of consensus"
6976..7665
/note="L1MA2 repeat: matches 4113..4798 of consensus"
8039..9213
/note="L1MA2 repeat: matches 2945..4113 of consensus"
9523..10437
/note="L1MA2 repeat: matches 2016..2945 of consensus"
10736..11494
/note="L1MA2 repeat: matches 1262..2016 of consensus"
11807..12251
/note="L1MCA repeat: matches 310..755 of consensus"
12259..12331
/note="HERVH21 repeat: matches 5..77 of consensus"
12332..12796
/note="L1R21A repeat: matches 1..490 of consensus"
16162..16371
/note="105 copies 2 mer tt 55% conserved"
17121..17424
/note="MER33 repeat: matches 1..322 of consensus"

repeat_region
18259..18290
/note="16 copies 2 mer ca 87% conserved"
18505..18672
/note="MER31A repeat: matches 1..181 of consensus"
19478..19814
/note="MER2 repeat: matches 1..345 of consensus"
21061..21101
/note="HERVH21 repeat: matches 42..85 of consensus"
21247..21781
/note="MLTIE repeat: matches 44..568 of consensus"
24465..24492
/note="Single clone region
Single clone region"
24531..24562
/note="Weak data
Weak data"
24563
/note="Single clone region
Single clone region"
24711..24865
/note="Region covered by subclones from a PAC DNA PCR
product only at 12x coverage.
Region covered by subclones from a PAC DNA PCR product only
at 12x coverage."
24758..25460
/note="CpG island"
/evidence="not_experimental"
24875..24940
/note="33 copies 2 mer cc 66% conserved"
26795..26926
/note="66 copies 2 mer ct 63% conserved"
27088..27133
/note="23 copies 2 mer tg 100% conserved"
27461..27777
/note="L1MB8 repeat: matches 5851..6171 of consensus"
27802..28232
/note="L1MC/D repeat: matches 5192..5403 of consensus"
28536..28930
/note="L1MC/D repeat: matches 4802..5192 of consensus"
29240..29341
/note="TIGGER1 repeat: matches 484..587 of consensus"
29603..30740
/note="TIGGER1 repeat: matches 587..1757 of consensus"
31027..31365
/note="TIGGER1 repeat: matches 1757..2105 of consensus"
31664..31691
/note="TIGGER1 repeat: matches 2105..2132 of consensus"
31809..32015
/note="TIGGER1 repeat: matches 2132..2246 of consensus"
32323..32389
/note="TIGGER1 repeat: matches 2246..2313 of consensus"
34353..34636
/note="MLT1-INTERNAL repeat: matches 607..889 of
consensus"
35149..35376
/note="MLT1-INTERNAL repeat: matches 1001..1233 of
consensus"
35523..35670
/note="MLTIC repeat: matches 8..159 of consensus"
BASE COUNT 9708 a 7094 c 6805 g 12107 t
ORIGIN

alignment_scores:

Quality: 471.50 Length: 129
Ratio: 4.100 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 70.543

alignment_block:

US-09-327-750D-33 x HS198P4/rev ..

Align seg 1/1 to reverse of: HS198P4 from: 1 to: 35714


```

1 MetGluSerLysAspGlnGlyAlaLysAsn...LeuAsnMetGluAsnAs 16
24358 ATGGAGTCCCAAGAGAACGAGCAGTAACAGTCTCAGCATGGAATAATGC 24309

16 pHisGlnLysLysGluGlnLysGluGlnLysProGlnAspThrIleLysA 33
24308 CAACCAAGAAATGAAGAAAGAG...CAAGTTGCTAATAAAG 24268

33 rgGluProValAlaIalProThrPheGluAlaGlyLysAsnCysAlaPro 49
24267 GGGAGCCCTTGGCCCTCCCT...TTGGATGCTGGTGAATACTGTGCTGCT 24221

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyAr 66
24220 AGAGAAATCGTAGGGGTTCCCGTTAGGCAGCCCATCTCAGTATAG 24171

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
24170 ATGGATATGATGATAGGCTTGGAGAACACAGCAGCAAGGATGAGAGAAG 24121

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
24120 AGAATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTG 24071

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
24070 AGGAAAGAGGATTTGAGTCATAGTCTGCGGAGCAGTACAGCTGACCCCT 24021

116 oHisHisAspHisAspGluPheCysLeuMetPro 128
24020 TCACCATGACCATCATGATGAGTTTTCCTTATGCC 23984

seq_name: gb_pr:AF251053

seq_documentation_block: 744 bp mRNA PRI 15-APR-2001
LOCUS AF251053
DEFINITION Homo sapiens x-linked protein mRNA, complete cds.
ACCESSION AF251053
VERSION AF251053.1 GI:13625167
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Mao,Y., Xie,Y., Zhou,Z., Zhao,W., Zhao,S., Wang,W., Huang,Y.,
Wang,S., Tang,R., Chen,X. and Wu,C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life
Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.
China
FEATURES
source Location/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
101..487
/codon_start=1
/product="x-linked protein"
/protein_id="AAK34943.1"
/db_xref="GI:13625168"
/translation="MESKEERALNNLIVENVNDKEDEKQVANKGPEPLALPNVS
EYCVPRGNRRFRVQPIQYRMDIMHRLGEPQARMRENMERIGEEVRLMKLREK
QLSHSLAVSTDPDPHDHDEFLMP"
BASE COUNT 222 a 134 c 208 g 180 t
ORIGIN

alignment_scores:
Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:
US-09-327-750D-33 x AF251053
Align seg 1/1 to: AF251053 from: 1 to: 744

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
|||||
101 ATGGAGTCCCAAGAGAACGAGCAGTAAACAGTCTCAGCATGCGTGGAAATGT 150

16 pHisGlnLysLysGluGlnLysGluGlnLysProGlnAspThrIleLysA 33
|||||
151 CAACCAAGAAATGAAGAAAGAGTGAAGAGGAGCAAGTTGCTAATAAAG 200

33 rgGluProValAlaIalProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||
201 GGGAGCCCTTGGCCCTACCT...TTGAATGTTAGTAATACTGTGCTGCT 247

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyAr 66
|||||
248 AGAGAAACCGTAGCGGTTCCCGTTAGGCAGCCCATCTCGCAGTATAG 297

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||
298 ATGGACATATAATGATAGGCTTGGAGAGCCACAGCAAGGATGAGAGAG 347

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||
348 AGAATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTG 397

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||
398 AGGAAAGAGCAGTTGAGTCATAGTCTGCGGAGCAGTACAGCTGATCCCTC 447

116 oHisHisAspHisAspGluPheCysLeuMetPro 128
|||||
448 TCACCATGACCATCATGATGAGTTTTCCTTATGCC 484

seq_name: gb_pat:AX078272
seq_documentation_block: 898 bp DNA PAT 22-FEB-2001
LOCUS AX078272
DEFINITION Sequence 76 from Patent WO0107471.
ACCESSION AX078272
VERSION AX078272.1 GI:13157963
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Huijman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzadeh,F., Yang,J., Lu,D.A., Baughm,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 76 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="Incyte ID No: 5664154CB1"
BASE COUNT 250 a 186 c 251 g 211 t
ORIGIN

alignment_scores:
Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:
US-09-327-750D-33 x AX078272
Align seg 1/1 to: AX078272 from: 1 to: 898

```


Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:

US-09-327-750d-33 x HSV870H8 ..

Align seg 1/1 to: HSV870H8 from: 1 to: 31321

```

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
|||||.....:|||||.....:|||||.....:|||||.....:
3508 ATGGAGTCCAAAGAGGAGGAGCGCTTAACAATCTCATCTCGGAAAATGT 3557
|||||.....:|||||.....:|||||.....:|||||.....:
16 phISGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysA 33
|||||.....:|||||.....:|||||.....:|||||.....:
3558 CAACAGGAAAATGATGAAAAAGATGAAAAGGAGCAAGTTCCTAATAAAG 3607
|||||.....:|||||.....:|||||.....:|||||.....:
33 rgGluProValAlaProThrPheGluAlaGlyLysAsnCyAlaPro 49
|||||.....:|||||.....:|||||.....:|||||.....:
3608 GGGAGCCCTTGCCCTACCT...TTGAATGTTAGTGAATACTGTGCCT 3654
|||||.....:|||||.....:|||||.....:|||||.....:
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||.....:|||||.....:|||||.....:|||||.....:
3655 AGAGAAACCGTAGCGGCTTCGCGTTAGGCAGCCCATCTGCAGTATAG 3704
|||||.....:|||||.....:|||||.....:|||||.....:
66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||.....:|||||.....:|||||.....:|||||.....:
3705 ATGGGACATATGCTAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGG 3754
|||||.....:|||||.....:|||||.....:|||||.....:
83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||.....:|||||.....:|||||.....:|||||.....:
3755 AGAATATGAAAGGATGCGGAGGAGGTGAGACAGCTGATGGAAAAGCTG 3804
|||||.....:|||||.....:|||||.....:|||||.....:
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||.....:|||||.....:|||||.....:|||||.....:
3805 AGGAAAAGCAGCTGAGTCATAGTCTCGGGCAGTCAGCACTGATCCCC 3854
|||||.....:|||||.....:|||||.....:|||||.....:
116 ohISHisAspHisHisAspGluPheCyLysLeuMetPro 128
|||||.....:|||||.....:|||||.....:|||||.....:
3855 TCACCATGACCATCAGTATGATGTTTGCCTTATGCC 3891
|||||.....:|||||.....:|||||.....:|||||.....:

```

seq_name: gb_pr:AL133348

seq_documentation_block: 40584 bp DNA PRI 06-MAR-2000

LOCUS AL133348 Human DNA sequence from clone RPI-79P11 on chromosome Xq21.32-22.1.
DEFINITION Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed X-linked protein 2), ESTs, STSS, GSSs and a
putative CpG island, complete sequence.

ACCESSION AL133348
VERSION AL133348.8 GI:7076386
KEYWORDS HTG; Bex2; CpG island.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 40584)

AUTHORS Wilson, S.

TITLE Direct Submission

JOURNAL Submitted (02-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonesrequest@sanger.ac.uk

COMMENT On Feb 24, 2000 this sequence version replaced gi:6997869.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at

http://www.sanger.ac.uk/HGP/chrX
RPI-79P11 is from the library RPI-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone
RPI-79P11. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

The true left end of clone LU0XNC01-105G4 is at 40485 in this
sequence. The true right end of clone LU0XNC01-177E8 is at 100 in
this sequence.

FEATURES

Source

Location/Qualifiers
1..40584
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q21.32-22.1"
/clone="RPI-79P11"
/clone_lib="RPI-1"

```

repeat_region 19..292  
/note="AluJo repeat: matches 3..275 of consensus"  
repeat_region 470..1284  
/note="L1MB5 repeat: matches 5318..6157 of consensus"  
repeat_region 1295..1898  
/note="L1PA15 repeat: matches 5556..6157 of consensus"  
repeat_region 1899..2345  
/note="L1PA15 repeat: matches 4922..5370 of consensus"  
repeat_region 2323..2698  
/note="L1MB5 repeat: matches 4960..5343 of consensus"  
repeat_region 2927..3065  
/note="MER74A repeat: matches 164..305 of consensus"  
repeat_region 3635..3955  
/note="L1PA4 repeat: matches 2..6142 of consensus"  
repeat_region 9866..9982  
/note="FLAM_C repeat: matches 6..124 of consensus"  
repeat_region 9989..10254  
/note="L1 repeat: matches 4027..4293 of consensus"  
repeat_region 10253..10852  
/note="L1M4 repeat: matches 4755..5333 of consensus"  
repeat_region 10853..11260  
/note="MUTB repeat: matches 3..390 of consensus"  
repeat_region 11261..11474  
/note="L1M4 repeat: matches 4538..4755 of consensus"  
repeat_region 11480..11563  
/note="MSTB repeat: matches 1..80 of consensus"  
repeat_region 11564..12475  
/note="L1 repeat: matches 3039..3971 of consensus"  
misc_feature 14463..14861  
/note="match: GSS: Em:AQ729215"  
repeat_region 15578..15854  
/note="AluJo repeat: matches 6..289 of consensus"  
misc_feature complement(16007..16556)  
/note="match: GSS: Em:AQ772810"  
repeat_region 16197..16569  
/note="L1M4 repeat: matches 4267..4627 of consensus"  
repeat_region 16329..16714  
/note="L1M4C repeat: matches 1578..1666 of consensus"  
repeat_region 16902..17133  
/note="L1M4 repeat: matches 0..242 of consensus"  
repeat_region 17536..17728  
/note="MIR repeat: matches 1..191 of consensus"  
repeat_region 17746..18323

```

```

repeat_region /note="L2 repeat: matches 1599. .2175 of consensus"
18706. .18755
repeat_region /note="L1M3 repeat: matches 6034. .6083 of consensus"
18756. .19326
repeat_region /note="MER66B repeat: matches 1. .486 of consensus"
19337. .21108
/note="MER66-internal repeat: matches 4808. .6676 of
consensus"
repeat_region 21109. .21163
/note="L1M2B repeat: matches 391. .445 of consensus"
repeat_region 21164. .21555
/note="MER66B repeat: matches 3. .391 of consensus"
repeat_region 21557. .21662
/note="L1M3 repeat: matches 6080. .6185 of consensus"
repeat_region 21663. .21761
/note="L1M repeat: matches 5299. .5392 of consensus"
repeat_region 21762. .21929
/note="L1M8 repeat: matches 5902. .6063 of consensus"
repeat_region 21930. .22212
/note="AluSg repeat: matches 3. .284 of consensus"
repeat_region 22213. .22321
/note="L1M8 repeat: matches 6063. .6173 of consensus"
repeat_region 22333. .22740
/note="L1M2 repeat: matches 5867. .6288 of consensus"
repeat_region 22744. .23094
/note="L1M2 repeat: matches 2. .342 of consensus"
repeat_region 23218. .23524
/note="AluSg repeat: matches 1. .305 of consensus"
repeat_region 23603. .23947
/note="L1M4 repeat: matches 4916. .5184 of consensus"
repeat_region 23948. .24218
/note="AluJo repeat: matches 1. .282 of consensus"
repeat_region 24219. .24510
/note="L1M4 repeat: matches 4601. .4916 of consensus"
repeat_region 24508. .24621
/note="L1M4 repeat: matches 4374. .4488 of consensus"
repeat_region 24650. .25077
/note="L1M4 repeat: matches 1. .466 of consensus"
repeat_region 25078. .25914
/note="L1M4 repeat: matches 3497. .4346 of consensus"
repeat_region 25920. .26741
/note="L1M4 repeat: matches 2202. .3065 of consensus"
repeat_region 26762. .26906
/note="L1P7 repeat: matches 6001. .6145 of consensus"
repeat_region 26907. .27135
/note="L1P5 repeat: matches 5674. .5902 of consensus"
repeat_region 27266. .27509
/note="L1M4 repeat: matches 2164. .2080 of consensus"
repeat_region 27557. .27769
/note="L1M3 repeat: matches 253. .460 of consensus"
repeat_region 28120. .28167
/note="L2 repeat: matches 2703. .2750 of consensus"
repeat_region 29043. .29349
/note="MER66B repeat: matches 1. .341 of consensus"
repeat_region 30389. .30452
/note="16 copies 4 mer tata.89% conserved"
polyA_site complement(30502)
mRNA /gene="dJ79p11.1"
complement(join(30502. .31137.31514. .31589.32016. .>32134))
/note="match: ESTs: Em:AA317587 Em:AA358632 Em:A1214048
Em:A1929470 Em:AA319216 Em:AA317625 Em:D81392 Em:AA317418
Em:AA365528 Em:CI5186 Em:A1291270 Em:AA160660 Em:AA612605
Em:AW160832 Em:AA384405 Em:AA160657 Em:AA081744
Em:A1291126 Em:CI5591 Em:AA384268 Em:A1929703 Em:A1302026
Em:W19547 Em:A1141727 Em:AA946933 Em:A1141583 Em:H61106
Em:W40581 Em:A1826617 Em:AA706545 Em:AA464771 Em:AA862327
Em:W48380 Em:AA641652 Em:AA384269 Em:AA527696 Em:W48778
Em:AA633989 Em:A1659486 Em:AW167301 Em:A1929428
Em:AW379411 Em:CI5948 Em:A1745669 Em:AW024587 Em:AW379207;
match: proteins: Tr:Q9Y516 Tr:Q9Y517"
/evidence="not_experimental"
/product="dJ79p11.1 (novel protein similar to mouse Bex2

```

```

(brain-expressed X-linked protein 2))"
complement(30502. .32134)
/gene="dJ79p11.1"
misc_feature 30507. .30988
/note="match: STS: Em:G24641; match: STS: Em:T86927"
misc_feature 30508. .30852
/note="match: STS: Em:R43117; match: STS: Em:G23964"
polyA_signal complement(30518. .30523)
/gene="dJ79p11.1"
misc_feature 30541. .30922
/note="match: STS: Em:H61107"
misc_feature 30543. .32262
/note="CpG island"
repeat_region 31647. .31730
/evidence="not_experimental"
/note="42 copies 2 mer cc 64% conserved"
misc_feature complement(32655. .33316)
/note="match: GSS: Em:AQ489303"
repeat_region 33027. .34514
/note="L1M2 repeat: matches 4631. .6155 of consensus"
repeat_region 34528. .34786
/note="L1M3A repeat: matches 4404. .4664 of consensus"
repeat_region 34787. .35085
/note="AluSg repeat: matches 1. .300 of consensus"
repeat_region 35086. .35412
alignment_scores:
Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217
alignment_block:
US-09-327-750D-33 x AL133348/rev
Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584
1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
31132 ATGGAGTCCAAGAGGACGCGTTAAACAATCTCATCTGGAATAATGT 31083
16 phiSGlnLysLysGluLysGluLysGluLysProGlnAspThrIleLysA 33
31082 CAACAGGAAATGATGATAAAGATGAAGACCAAGTTGCTTAATAAAG 31033
33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
31032 GGGAGCCCTTGGCCCTAGCT...TTGAATGTTAGTAATACTGTGCT 30986
50 ATGGTGLVATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
30985 AGAGGAAACCGTAGCGGCTCCGCTTAGCGAGCCCATCTGCAGTATAG 30936
66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluC 83
30935 ATGGGACATATATGTCATAGCTTGGAGAGCCACAGGCAAGGATGAGAG 30886
83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
30885 AGAATATGAAGAGGATGGGAGGAGGTGACACAGCTGATGGAAGAGCTG 30836
100 ATGGTGLVATGATGATGATGATGATGATGATGATGATGATGATGATGAT 116
30835 AGGGAAACGAGTTCAGTATAGTTTCGGGCGAGTGCAGTATGATCCCC 30786
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
30785 TCACCATGACCATCAGTATGATGATGATGATGATGATGATGATGATGAT 30749
seq_name: gb_hgt:HSU80B1
seq_documentation_block:
LOCUS HSU80B1 41029 bp DNA HTG
DEFINITION Homo sapiens chromosome X clone LL0XNC01-80B1 map q22.1-22.3, ***

```


Align seg 1/1 to: AF237783 from: 1 to: 791

```

1 MetGluSerLysAspGlnGlyAlaLysAsp...LeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172 ATGGAGTCCAAAGAACTAGCAGTAACAGCTCAGCATGGAATAATGC 221
16 phiGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 CAACCAAGAAATGAAGAAAGAG.....CAAGTTGCTAATAAAG 262
33 rgGluProValAlaProThrPheGluAlaClyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 GGGAGCCCTGGCCCTCCCT...TTGATGCTGGTGAATACGTGTGCT 309
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
310 AGAGGAATCGTAGCGGTTCCGCGTTAGGCAGCCCATCTCGAGTATAG 359
66 gTspAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 ATGGATATGATGCTAGGCTTGGAACCCACAGGAGGATGAGAGAG 409
83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
410 AGAATATGGAAGGATTTGGGAGGGGTGAGACAGCTGATGGAAGCTG 459
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
460 AGGAAAGAGGTTGAGTATAGTCTGCGGGCAGTCAGCACTGACCCCC 509
116 oHisHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
510 TCACCATGACCATCATGATGAGTTTTCCTTATGCC 546

```

seq_name: gb_sts:G24641

```

seq_documentation_block: 504 bp DNA STS 31-MAY-1996
DEFINITION human STS WI-11354, sequence tagged site.
ACCESSION G24641
VERSION G24641.1 GI:1344967
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human STS derived from sequences in dbEST and the Unigene
collection.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hudson.T.
Whitehead Institute/MIT Center for Genome Research; Physically
Mapped STSs
Unpublished (1995)

```

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: TTTTGTGCTTTTAAAGTGC
 Primer B: TTACGTGGGTCTCCTATTACCA
 STS size: 150
 PCR Profile:

Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng

Primer: each 5 pM
 dNTPs: each 4 nM
 Taq Polymerase: 0.025 units/ul
 Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 9.3

Derived from dbEST (genbank accession #86927).

```

FEATURES             Location/Qualifiers
     source            1..504
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
     STS
     primer_bind       1..150
     primer_bind       1..22
     primer_bind       complement(129..150)
     BASE COUNT       134 a 118 c 101 g 144 t 7 others
     ORIGIN

```

alignment_scores:
 Quality: 311.00 Length: 85
 Ratio: 4.260 Gaps: 3
 Percent Similarity: 85.882 Percent Identity: 72.941

alignment_block:

US-09-327-750D-33 x G24641/rev ..

Align seg 1/1 to reverse of: G24641 from: 1 to: 504

```

47 CysAlaProArgGly.GlyArgArgArgPhe.ArgValArgGlnProIle 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
503 TGTCCTCCCTANAGGAAACCGTAGCGGTTCCCTTGGGAGGAGGAGGAG 454
63 SerHisTyrArgTspAspLeuMetHisArgValGlyGlu.ProGlnGlyA 79
:|||||:|||||:|||||:|||||:|||||:|||||:
453 CCGCAGTANAGTGGGATATGATGATGATGATGATGATGATGATGATGAT 404
79 rgMetArgGluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeu 95
|||||:|||||:|||||:|||||:|||||:|||||:
403 GGATGAGAGAGAGATATGGAAGGATGCGGAGGAGGAGGAGGAGGAGGAG 354
96 MetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSe 112
|||||:|||||:|||||:|||||:|||||:|||||:
353 ATGAAAAGCTGAGGAGAAAGCAGTTGAGTCANAGTCTCGGGCAGTCAG 304
112 rThrAspProHisHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:
303 CACTGACCCCTCACCATGACCATCATGATGAGTTTTCCTTATGCC 255

```

/cgn2_6/ptodata/2/1na/PCTUS_COMB.seq.PCT-US96-00419-6 +	68.50	119.81	50.46
/cgn2_6/ptodata/2/1na/6B_COMB.seq.US-08-847-065-24 +	68.50	119.22	54.45
/cgn2_6/ptodata/2/1na/6B_COMB.seq.US-09-308-022-4 +	68.50	117.13	71.21
/cgn2_6/ptodata/2/1na/6A_COMB.seq.US-09-128-155-17 -	68.50	75.66	1.46+04
/cgn2_6/ptodata/2/1na/6A_COMB.seq.US-09-074-912-3 +	68.00	125.24	25.17
/cgn2_6/ptodata/2/1na/6A_COMB.seq.US-09-171-151-3 +	68.00	125.24	25.17

seq_name: /cqn2_6/ptodata/2/lna/6B_COMB.seq:US-08-467-504-1

```
seq_documentation_block:
    : Sequence 1, Application US/08467504
    : Patent No. 6211430
    : GENERAL INFORMATION:
    : APPLICANT: John, Maliyakal E.
    : TITLE OF INVENTION: FLIATE PROMOTER
    : NUMBER OF SEQUENCES: 16
    : CORRESPONDENCE ADDRESS:
    : ADDRESSEE: Quarles & Brady
    : STREET: 411 East Wisconsin Avenue
    : CITY: Milwaukee
    : STATE: WI
    : COUNTRY: U.S.A.
    : ZIP: 53202-4497
    : COMPUTER READABLE FORM:
    : MEDIUM TYPE: Floppy disk
    : COMPUTER: IBM PC compatible
    : OPERATING SYSTEM: PC-DOS/MS-DOS
    : SOFTWARE: Patent In Release #1.0, Version #1.25
    : CURRENT APPLICATION DATA:
    : APPLICATION NUMBER: US/08467504
```

```

//
// CLASSIFICATION: 800
//
// ATTORNEY/AGENT INFORMATION:
//
// NAME: Baker, Jean C.
//
// REGISTRATION NUMBER: 35,433
// REFERENCE/DOCKET NUMBER: 110229.91152.
//
// TELECOMMUNICATION INFORMATION:
//
// TELEPHONE: (414) 277-5709
// TELEFAX: (414) 271-3552
//
// INFORMATION FOR SEQ ID NO. 1:
//
// SEQUENCE CHARACTERISTICS:
//
// LENGTH: 974 base pairs
//
// TYPE: nucleic acid
//
// STRANDEDNESS: double
//
// TOPOLOGY: linear
//
// MOLECULE TYPE: cDNA
//
// US-08-467-504-1
//
//
// alignment_scores:
//
// Quality: 74.50 Length: 92
// Ratio: 1.620 Gaps: 3
// Percent Similarity: 50.000 Percent Identity: 27.174
//
// alignment_block:
//
// US-09-327-750D-34 x US-08-467-504-1 ..
//
// Align seg 1/1 to: US-08-467-504-1 from: 1 to: 974
//
//
// 15 GluLysAspLysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGl 31
// |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
// 364 GAAAGGAAGATTAAGTAAAGTAAAGCAACATAAAGATGAAGAGTGCACGAGGTC 413
//
//
// 31 uGluGluSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyA 48
// |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
// 414 ACAGCAATCCAAAGAGCAGCACGAGTACGAGAAAGAAAACCCGAT.... 459
//
//
// 48 snValArgArgLysValArgLeuValProAsnPheLeuTrpAlaIle 64
// |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
// 460 .....TTCCCAAA.....TGGGAAAG 477
//
//
// 65 ProAsnArgHisValAspHisSer.....GluGl 74
// |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
// 478 CCTTAAGGGCCAGGAACATAAAGAGCCGGAATATCCGAAATACCTCAGTG 527

```



```
;
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 326..6277
; US-08-836-325-9

alignment_scores:
  Quality: 72.50 Length: 58
  Ratio: 1.812 Gaps: 4
  Percent Similarity: 68.966 Percent Identity: 36.207

alignment_block:
  US-09-327-750D-34 x US-08-836-325-9/rev ..
  Align seg 1/1 to reverse of: US-08-836-325-9 from: 1 to: 6452

325 LysAlaSerLysGlnSerGluGluGlu.....SerHisHisLeuG1 38
|||||:|||||:|||||:|||||:|||||:
3220 AAGAGCCAGAAACAGGTTCCAGGACCAAGGTTCCCAATCACCATGACCA 3171

38 uGluValGluAsnLysLysProGlyGlyAsnValArgArgLysValArgA 55
|||||:|||||:|||||:|||||:|||||:
3170 TCATGTAAACAAATPAGGCACATGGT.....CTGCCCGCGACCTC...C 3130

55 rgLeuValProAsnPhLeuTptAlaLalleProAsnArgHisValAspHis 71
|||||:|||||:|||||:|||||:|||||:
3129 ATCAGTCCCACTGCTCTCTATCCACTCCCCACACAGCAGCTCGGAACAC 3080

72 SerGluGlyGly...GluGluVal 78
|||||:|||||:|||||:|||||:
3079 GATGAGGAGGAGGTGGGAAGATC 3056

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-892-770-1

seq_documentation_block:
; Sequence 1, Application US/08892770
; Patent No. 5962670
; GENERAL INFORMATION:
; APPLICANT: Walling, Linda L.
; APPLICANT: Pautot, Veronique
; APPLICANT: Gu, Yong-Qiang
; APPLICANT: Chao, Wun Shaw
; TITLE OF INVENTION: Improved Promoters for Enhancing Plant
; TITLE OF INVENTION: Productivity
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,770
; FILING DATE: 15-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Quine, Jonathan A.
; REGISTRATION NUMBER: P-41,261
; REFERENCE/DOCKET NUMBER: 023070-072100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 935 base pairs

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY:
; LOCATION: 1..935
; OTHER INFORMATION: /note= "tomato acidic leucine
; OTHER INFORMATION: aminopeptidase 1 (Lapal) promoter
; OTHER INFORMATION: region"
; US-08-892-770-1

alignment_scores:
  Quality: 72.00 Length: 60
  Ratio: 2.250 Gaps: 2
  Percent Similarity: 53.333 Percent Identity: 28.333

alignment_block:
  US-09-327-750D-34 x US-08-892-770-1 ..
  Align seg 1/1 to: US-08-892-770-1 from: 1 to: 935

62 TrpalalleProAsnArgHisValAspHisSerGluGlyGlyGluGluVa 78
|||||:|||||:|||||:|||||:|||||:
438 TGGTGTGTGTGAATGAGATTATCGATCATATAAGAGAGGGGAGAGAT 487

78 lGlyArgPheValGlyGlnValMet.GluAlaLysArgHis..... 91
|||||:|||||:|||||:|||||:
488 GAACAGTTTTTAAATAAATTTTGGCCCCAGAAAAAACATTCCACGCGT 537

92 .....Serly 93
|||||

538 GACTTTAGTCATTACACTCCTTTGCCATGACAGAGAAATCGTCAAA 587
|||||
93 sGluGlnGlnMetArgProTyrThrArg 102
|||||
588 AAATGACACATGAGACCTTACATAAGG 615

seq_name: /cgn2_5/ptodata/2/ina/6A_COMB.seq:US-08-726-214-5

seq_documentation_block:
; Sequence 5, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gliman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,214
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450
```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA

alignment_scores:
 Quality: 71.00 Length: 100
 Ratio: 1.203 Gaps: 3
 Percent Similarity: 59.000 Percent Identity: 20.000

alignment_block:

US-09-327-750D-34 x 5180810-2 ..

Align seg 1/1 to: 5180810-2 from: 1 to: 1784

```

3 SerLysValLysGlnValIleLeuAspLeuThrValGluLysAspLysLys 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
505 GCTAACTGTTGAGGTGTTGAAACCAAGCTTGGAAACGAA..... 549
19 sAsnLysGlyGlyLysLysSerLysGlnSerGlnSerGluGluSerHis 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
550 .....AAACTCAGAGTGCAGATGAGGAGATAGAGAA 583
36 IsLeuGluGluValGluAsnLysLysProGlyGlyAsnValArgLys 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
584 ATTAGCAAACTAGCAAGAAATCAAGGA.....AAG 618
53 ValArgLysLeuValProAsnPheLeuTrpAlaIleProAsnArgHisVa 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
619 CTCGAAAAATGGAGCTTGACTATCTCAAAAAATTAGATCAGGACACAA 669
69 LasPHisSerGlyGlyGlyGluValGlyArgPheValGlyGlnValm 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
669 AGACGACCAAAAGACACACAA.....C 694
86 etGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyrThrArg 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
695 AAGAGAGCGCAAAAAAATCAAGAACATTAGAAGCTAAATACCAACGA 744

```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-841-483-3

seq_documentation_block:

; Sequence 3, Application US/08841483B

; Patent No. 5976875

; GENERAL INFORMATION:

; APPLICANT: Prescott, Steven M.

; APPLICANT: Bunting, Michaeline

; APPLICANT: Tang, Wen

; APPLICANT: Topham, Matthew

; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and

; FILE REFERENCE: 2037.2.1a

; CURRENT APPLICATION NUMBER: US/08/841.483B

; EARLIER FILING DATE: 1997-04-22

; EARLIER FILING DATE: 1996-04-22

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 3490

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (89)..(2875)

US-08-841-483-3

alignment_scores:

Quality: 71.00 Length: 82
 Ratio: 1.651 Gaps: 4
 Percent Similarity: 52.439 Percent Identity: 30.488

alignment_block:

US-09-327-750D-34 x US-08-841-483-3 ..

Align seg 1/1 to: US-08-841-483-3 from: 1 to: 3490

```

13 ThrValGluLysAspLysLysAsnLysLysGlyGly.....Lysal 26
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
848 ACTCTGAAAGCAAGCAAGAGAGAGAGGCGATCTTCAAGAGGAATC 897
26 aserLysGlnSerGluGluSerHis..... 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
898 CAGCAAGAAAGGCGCTGAGGAGGCGCGCTGGAGACCCCTTCATCATCAGGC 947
36 .....HisLeuGluGluValGluAsnLysLys 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
948 CCACCCCTCCCGCTCATCAAGCCCTGCTGCTGTTGTGAACCCCAAG 997
45 ProGlyGlyAsnValArgLysValArgLeuValProAsnPheLe 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
998 AGTGGGGGACACAGGGTGCAGAGATC.....ATCCAGTCTTCTCT 1038
61 uTrpAlaIleProAsnArgHisVal....AspHisSerGluGlyGly 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1039 CTGGTATCTCAATCCCGACAAAGTCTTCGACCTGAGCCAGGAGGG 1084

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-382-911-3

seq_documentation_block:

; Sequence 3, Application US/09382911

; Patent No. 6221658

; GENERAL INFORMATION:

; APPLICANT: Prescott, Steven M.

; APPLICANT: Bunting, Michaeline

; APPLICANT: Tang, Wen

; APPLICANT: Topham, Matthew

; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and

; FILE REFERENCE: 2037.2.1a

; CURRENT APPLICATION NUMBER: US/09/382.911

; PRIOR FILING DATE: 1999-08-25

; PRIOR FILING DATE: 1997-04-22

; PRIOR APPLICATION NUMBER: 60/016.210

; PRIOR FILING DATE: 1996-04-22

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 3490

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (89)..(2875)

US-09-382-911-3

alignment_scores:

Quality: 71.00 Length: 82
 Ratio: 1.651 Gaps: 4
 Percent Similarity: 52.439 Percent Identity: 30.488

alignment_block:

US-09-327-750D-34 x US-09-382-911-3 ..

Align seg 1/1 to: US-09-382-911-3 from: 1 to: 3490

```

13 ThrValGluLysAspLysLysAsnLysLysGlyGly.....Lysal 26
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
848 ACTCTGAAAGCAAGCAAGAGAGAGGCGATCTTCAAGAGGAATC 897
26 aserLysGlnSerGluGluSerHis..... 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
898 CAGCAAGAAAGGCGCTGAGGAGGCGCGCTGGAGACCCCTTCATCATCAGGC 947
36 .....HisLeuGluGluValGluAsnLysLys 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
948 CCACCCCTCCCGCTCATCAAGCCCTGCTGCTGTTGTGAACCCCAAG 997

```



```

; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,043B
; FILING DATE: 19911212
; CLASSIFICATION: 434
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; FILING DATE: 9-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-May-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5342774man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 253.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4698 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-07-807-043B-5

alignment_scores:
  Quality: 71.00      Length: 93
  Ratio: 1.365      Gaps: 4
  Percent Similarity: 55.914      Percent Identity: 27.957

alignment_block:
US-09-327-750D-34 x US-07-807-043B-5/rev ..
Align seg 1/1 to reverse of: US-07-807-043B-5 from: 1 to: 4698

16 LysAspLysLysAsnLysGlyLysAlaSer.LysGlnSerGluG 32
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2134 AGGGACAGGAAGGCAAGAGGGGGGAGCCGCGAGGGGGGAGC 2085
:::|||||:|||||:|||||:|||||:|||||:|||||:

32 luGluSerHisHisLeuGluGluValGluAsnLys.....Lys 44
:::|||||:|||||:|||||:|||||:|||||:|||||:
2084 AGGAAGGCACACAGAGGGAAGCGGAGGGGGGGGATCTTGAAAA 2035
:::|||||:|||||:|||||:|||||:|||||:|||||:

45 ProGlyGlyAsnValArgLysValArgLysValProAsnPheLe 61
|||||:|||||:|||||:|||||:|||||:|||||:
2034 CAGGTGCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1990
|||||:|||||:|||||:|||||:|||||:|||||:

61 uTrpAlaIleProAsnArgHisValAspHisSerGluGlyGluGlu 78
|||||:|||||:|||||:|||||:|||||:|||||:
1989 .....CCTAACCAACCAACCAACCAACCAACCAACCAAC 1950
|||||:|||||:|||||:|||||:|||||:|||||:

78 alGly.....ArgPheValGlyGlnValMetGluAlaLys 89
|||||:|||||:|||||:|||||:|||||:|||||:
1949 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
|||||:|||||:|||||:|||||:|||||:|||||:

90 ArgHisSerLysGluGlnMetArg 98
:::|||||:|||||:|||||:|||||:
1899 AAGGCTGCTAAAGTCGACGAGGAGGAGGAGGAGGAGGAGG 1873

seq_name: /cnp2_6/ptodata/2/ina/5A_COMB.seq:US-08-299-849B-5

seq_documentation_block:
; Sequence 5, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
; APPLICANT: De Plaeen, Etienne; Boon-Falleur, Thierry;
; APPLICANT: Leth, Bernard; Szikora, Jean-Pierre; De Smet, Charles;
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
; APPLICATION OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felie & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,849B
; FILING DATE: 1-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/037,230
; FILING DATE: 26-MARCH-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04354
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/807,043
; FILING DATE: 12-DECEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; FILING DATE: 9-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-May-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5612201man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5355
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4698 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-299-849B-5

alignment_scores:
  Quality: 71.00      Length: 93
  Ratio: 1.365      Gaps: 4
  Percent Similarity: 55.914      Percent Identity: 27.957

alignment_block:
US-09-327-750D-34 x US-08-299-849B-5/rev ..
Align seg 1/1 to reverse of: US-08-299-849B-5 from: 1 to: 4698

16 LysAspLysLysAsnLysGlyLysAlaSer.LysGlnSerGluG 32
:::|||||:|||||:|||||:|||||:|||||:|||||:
2134 AGGGACAGGAAGGCAAGAGGGGGGAGCCGAGGGGGGAGGAGC 2085
:::|||||:|||||:|||||:|||||:|||||:|||||:

32 luGluSerHisHisLeuGluGluValGluAsnLys.....Lys 44
:::|||||:|||||:|||||:|||||:|||||:|||||:
2084 AGGAAGGCACACAGAGGGAAGCGGAGGGGGGGGATCTTGAAAA 2035
:::|||||:|||||:|||||:|||||:|||||:|||||:

45 ProGlyGlyAsnValArgLysValArgLysValProAsnPheLe 61
|||||:|||||:|||||:|||||:|||||:|||||:
2034 CAGGTGCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1990
|||||:|||||:|||||:|||||:|||||:|||||:

61 uTrpAlaIleProAsnArgHisValAspHisSerGluGlyGluGlu 78
|||||:|||||:|||||:|||||:|||||:|||||:
1989 .....CCTAACCAACCAACCAACCAACCAACCAACCAAC 1950
|||||:|||||:|||||:|||||:|||||:|||||:

78 alGly.....ArgPheValGlyGlnValMetGluAlaLys 89
|||||:|||||:|||||:|||||:|||||:|||||:
1949 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
|||||:|||||:|||||:|||||:|||||:|||||:

90 ArgHisSerLysGluGlnMetArg 98
:::|||||:|||||:|||||:|||||:
1899 AAGGCTGCTAAAGTCGACGAGGAGGAGGAGGAGGAGGAGGAG 1873

seq_name: /cnp2_6/ptodata/2/ina/5A_COMB.seq:US-08-299-849B-5

seq_documentation_block:
; Sequence 5, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
```

Tue Mar 12 09:01:44 2002

```
45 ProGlyGlyAsnValArgArgLysValArgArgLeuValProAsnPheLe 61
|||:||||: |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
2034 CAAGGTGCAAAAAAAAAAAAAAAAAAGCAACCAACCAACCAACCAACAA..... 1990

61 uTrpAlaIleProAsnArgHisValAspHisSerGluGlyGluGluV 78
|||||:||||: |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
1989 .....CCAAACAACACAAAAAGGTGAGCTGGAGGTGGAGGGGAGC 1950

78 alGly.....ArgPheValGlyGlnValMetGluAlaLys 89
|||:||||: |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
1949 CAGGAGGAAGGCACAAAAAGGGGACGCGGCAGAAAAAGCAGGAGGAGAGGT 1900

90 ArgHisSerLysGluGlnGlnMetArg 98
|||:||||: |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
1899 AAGCTGCTAAAGTCGAGCAAGAGG 1873
```

OM of: US-09-327-750D-34 to: N_Geneseq_1101.* out_format : pfs
 Date: Mar 11, 2002 3:42 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-framet_B2n.model -DEV-xlh
 -Q/cn2_1/USPFO_Spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
 -DB-N_Geneseq_1101 -QFWT-fastap -SUFFIX-rng -GAPOP-12.000
 -GAPEXT-4.000 -MINMATCH-0.100 -LOOPCL-0.000 -LOOPEXT-0.000
 -GAPOP-4.500 -QGAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500
 -FGAPOP-6.000 -FGAPEXT-7.000 -XGAPOP-10.000 -YGAPEXT-0.500
 -DELOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-blosum62
 -TRANS-human40.cdi -LIST-45 -DOCALIGN-200 -THR_SCORE-pct
 -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFWT-pfs
 -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
 -USER-US09327750 -CGN1_1330 -NCPU-6 -ICPU-3 -LONGLOG
 -DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -NO_XLPXY -WAIT -THREADS-1

Search information block:

Query: US-09-327-750D-34
 Query length: 118
 Database: N_Geneseq_1101.*
 Database sequences: 930621
 Database length: 428662619
 Search time (sec): 355.560000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517 +		372.00	800.02	865	1.7e-36
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750 +		372.00	796.41	1229	2.8e-36
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH35548 +		372.00	795.34	1364	3.2e-36
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23529 +		171.50	358.04	891	7.2e-12
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748 +		170.00	354.44	917	1.1e-11
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528 +		144.00	299.94	700	1.2e-08
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01004 +		143.00	301.51	485	1.0e-08
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01005 +		136.00	285.14	532	8.3e-08
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15811 +		111.00	226.31	767	0.0002
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15811 +		111.00	225.10	862	0.0002
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH59611 +		111.00	224.68	898	0.0002
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143 +		109.50	222.67	792	0.0003
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03880 +		103.00	210.20	662	0.0012
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03879 +		103.00	209.65	698	0.0013
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03867 +		99.00	198.72	858	0.0034
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH0889 +		98.50	199.83	692	0.0047
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01003 +		97.50	203.23	401	0.0030
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH23519 +		92.50	143.81	44453	6.20
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH20201 +		89.50	189.45	276	0.0178
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15401 +		89.50	189.45	276	0.0178
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH05905 +		89.50	189.45	276	0.0178
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13308 +		87.50	180.85	415	0.0336
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13461 +		87.50	180.85	415	0.0336
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03209 +		87.50	180.85	415	0.0336
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH23789 +		83.50	158.39	9561	0.9561
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH10225 +		83.50	114.98	250.27	106746
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH06235 +		81.00	150.02	2.80	2.80
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH15484 +		80.50	159.18	0.8631	0.8631
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH28355 +		80.00	141.62	262.02	52216
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH46153 +		78.00	137.19	14.49	3786
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH7225 +		78.00	136.98	14.89	3866
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH08642 +		78.00	135.42	15.19	3925
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH06237 +		78.00	135.42	18.19	4500
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH06236 +		78.00	135.31	18.45	4548
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH07222 +		77.50	131.37	30.58	5994
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH39709 +		77.00	150.83	2.52	811
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH22774 +		77.00	147.68	3.77	1102
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH121493 +		76.00	156.05	1.29	394
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH16784 +		76.00	156.05	1.29	394
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH01790 +		76.00	156.05	1.29	394

/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH54129 + 76.00 133.47 23.35
 /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH93013 - 75.50 149.04 3.17
 /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH73887 - 75.50 142.27 7.55
 /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH72409 + 75.00 142.42 7.41
 /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH00010 - 75.00 133.35 23.71

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA; 865 BP.
 XX
 AC AAH03517;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:352.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 (HELI-) HELIX RES INST.
 PA
 PI Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sujiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 WPI: 2001-318749/34.
 XX
 Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs -
 Claim 1; SEQ ID 352; 2537pp + CD ROM; English.
 The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesizing polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any special methods. AAH03166 to AAH13628 and
 AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.
 Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:
 Quality: 372.00 Length: 120
 Ratio: 3.758 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:
 US-09-327-750D-34 x AAH03517 ..

Align seg 1/1 to: AAH03517 from: 1 to: 865

```

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
||||| ||||| ... ..: |||||: |||||:
191 ATGGAGTCCAAAGAGGAGTACTAGCGCAACAATCTCAACGGGAAATGC 240
17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
241 CCAACAAGAAAACGAAGAGGGGAGCAGGCCGCCACGCAAGATGAAGAAG 290
33 LuSerHisHisLeuGluValGluAsnLysLysProGlyGlyAsnVal 49
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
291 AATCCCGCCATTGGGAGGGGGTGAAGCCAGAGCCTGGAGGAATATC 340
50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
341 AGCGGGGGCGAGTGTAGCGACTTGTCCCTAATTTTCGATGGCCATACC 390
65 OAsnArgHisValAspHisSerGluGlyGlyGluValGlyArgPheV 82
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 TAATAGGCATATTGAGCACAATGAAGCGAGAGATGATGAGAAAGGTTG 440
82 aLcGlyClnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArg 98
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 TAGGCAGATGATGGAATCAAGAGAAGACTAGGGACACAGCAGATGAGG 490
99 ProTyThrArgPheArgThrProGluProAspAsnHisTyRAspPheCy 115
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 CACTATATCGCTTCCAACTCTGTGACCTGACCAACCATTTATGACTTTG 540
115 sleullePro 118
||||| |||||
541 CCTCATACCT 550

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH13750

seq_documentation_block:
 ID: AAH13750 standard; CDNA; 1229 BP.

XX AC AAH13750;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human cDNA sequence SEQ ID NO:10656.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI: 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX SQ Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:
 Quality: 372.00 Length: 120
 Ratio: 3.758 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:
 US-09-327-750D-34 x AAH13750 ..

Align seg 1/1 to: AAH13750 from: 1 to: 1229

```

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
||||| ||||| ... ..: |||||: |||||: |||||:
191 ATGGAGTCCAAAGAGGAGTACTAGCGCAACAATCTCAACGGGAAATGC 240
17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
241 CCAACAAGAAAACGAAGAGGGGAGCAGGCCGCCACGCAAGATGAAGAAG 290
33 LuSerHisHisLeuGluValGluAsnLysLysProGlyGlyAsnVal 49
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
291 AATCCCGCCATTGGGAGGGGGTGAAGCCAGAGCCTGGAGGAATATC 340
50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
341 AGCGGGGGCGAGTGTAGCGACTTGTCCCTAATTTTCGATGGCCATACC 390
65 OAsnArgHisValAspHisSerGluGlyGlyGluValGlyArgPheV 82
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 TAATAGGCATATTGAGCACAATGAAGCGAGAGATGATGAGAAAGGTTG 440
82 aLcGlyClnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArg 98
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 TAGGCAGATGATGGAATCAAGAGAAGACTAGGGACACACAGCAGATGAGG 490
99 ProTyThrArgPheArgThrProGluProAspAsnHisTyRAspPheCy 115
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 CACTATATCGCTTCCAACTCTGTGACCTGACCAACCATTTATGACTTTG 540

```

115 sleuillePro 118
|||||
541 CCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.AAC85548

seq_documentation_block:

ID AAC85548 standard; cDNA; 1364 BP.

XX AAC85548;

XX 04-JUN-2001 (first entry)

XX cDNA encoding CDIFF-4, Incyte ID No. 1990956CB1.

XX Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
KW cell proliferation; Alzheimer's disease; schizophrenic disorder;
KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 250..612
XX /*tag= a
XX /product= "CDIFF-4"

XX WO200119860-A2.

XX 22-MAR-2001.

XX 14-SEP-2000; 2000WO-US25435.

XX 15-SEP-1999; 99US-0154140.

XX 06-DEC-1999; 99US-0169155.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;

XX WPI: 2001-211447/21.

XX P-PSDB: AAB47126.

XX Isolated polypeptides and polynucleotides involved in cell
PT differentiation are used for treatment, prevention and diagnosis of
PT cell proliferative, developmental and neurological disorders e.g.
PT cancer and Alzheimer's disease -

XX Claim 5; Page 121; 137pp; English.

XX The sequences given in AAC85545-72 encode human polypeptides involved
CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
CC these are used to treat a disease or condition associated with
CC decreased expression of functional CDIFF. An antagonist of CDIFF is
CC used to treat a disease or condition associated with over expression
CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
CC prevention and diagnosis of cell proliferative, developmental and
CC neurological disorders, such as Alzheimer's disease, schizophrenic
CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
CC REX-3. This sequence maps to chromosome 1 within the interval from
CC 152.2 to 157.4 centiMorgans, to chromosome 3 within the interval from
CC 157.4 to 158.0 centiMorgans, and to the X chromosome within the interval
CC from 104.9 to 150.3 centiMorgans.

XX Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:

Quality: 372.00 Length: 120
Ratio: 3.758 Gaps: 2
Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:

US-09-327-750D-34 x AAC85548

Align seq 1/1 to: AAC85548 from: 1 to: 1364

1 MetAlaSerLysValLysGlnValLleLeuAspLeuThrValGluLysAs 17
250 ATGGAGTCCAAAGAGAGAACTAGCGGCAACAAATCTCAACGGGGAATGC 299
17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
300 CCAACAAGAAACGAAGGAGGAGGAGCGCCGCCACGAGAATGAAGAAG 349
33 LuSerHisLysLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49
350 AATCCCGCCCATTTGGAGGGGTGAAGCCAGACGCTGGAGGAATATC 399
50 ArgArg...LysValArgArgLeuValProAsnPheLeuTlpAlaIlePr 65
400 AGCGGGGCGAGTTAGGCGACTTGCTCTAATTTTCGATGGCCATACC 449
65 OAsnArgHisValAspHisSerGluGlyGlyClyClyValGlyArgPheV 82
450 TAATAGGCATATTGAGCAATGAAGAGAGAGATGATGATGAAGAGTTG 499
82 alGlyGlnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArg 98
500 TAGGCAGATGATGGAATCAAGAGAGAGAGATGAGGAAACAGCAGATGAGG 549
99 ProTyrThrArgPheArgThrProGluProAsnHisTyrAspPheCy 115
550 CACTATATGCGCTTCCAACTCCTGAACCTGACACCAACATTATGACTTTG 599
115 sleuillePro 118
600 CCTCATACCT 609

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.AAF23529

seq_documentation_block:

ID AAF23529 standard; DNA; 891 BP.

XX AAF23529;

XX 22-MAR-2001 (first entry)

XX Human NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.

XX Homo sapiens.

XX WO200075278-A2.

XX 14-DEC-2000.

XX 07-JUN-2000; 2000WO-US15621.

XX 07-JUN-1999; 99US-0327750.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Sato T;

XX WPI: 2001-061707/07.

XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT the gene encoding NADE, useful for modulating the activity of p75NTR
PT and for detecting neurodegenerative diseases -

XX Disclosure: Fig 1; 134pp; English.

486 CATATGGAATATTCATGGAGGATGAGAGAAATCAGAGAAACTTAG 535
93 sGluGlnMetArgProTyrThrArgPheArgThrProGluProAspA 110
|||||
536 GGAGCTCAGTTGAGGAATGTCTGCTATCTTATGGGGAGCTCTCTA 585
110 snHisTyrAsp.....PheCysLeuIlePro 118
|||||
586 ATCACCATGACCATCATGATGAATTTGCCCTATGCCT 623

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF23528

seq_documentation_block:

ID AAF23528 standard; DNA; 700 BP.

AC AAF23528;

DT 22-MAR-2001 (first entry)

DE Mouse NADE DNA.

KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.

OS Mus sp.

PN W0200075278-A2.

PD 14-DEC-2000.

PF 07-JUN-2000; 2000WO-US15621.

PR 07-JUN-1999; 99US-0327750.

PA (UYCO) UNIV COLUMBIA NEW YORK.

PI Sato T;

DR WPI; 2001-061707/07.

PT New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT the gene encoding NADE, useful for modulating the activity of p75NTR
PT and for detecting neurodegenerative diseases -

PS Claim 12; Fig 1; 134pp; English.

CC The present invention relates to a purified polypeptide capable of
CC binding neurotrophin receptor (p75-NTR). The invention is useful for
CC NGF-induced apoptosis, which plays an important role in neurogenetic
CC diseases. The peptide of the invention and p75NTR are useful for
CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.

SQ Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

alignment_scores:

Quality: 144.00 Length: 129
Ratio: 2.000 Gaps: 3
Percent Similarity: 55.814 Percent Identity: 28.682

alignment_block:

US-09-327-750D-34 x AAF23528 ..

Align seg 1/1 to: AAF23528 from: 1 to: 700

18 LysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlucose 34

162 ANACAAATCTCATCTGCGCATGTCACAGGAACAGAGAGACTGA 211

34 rHisHisLeuGluGluValGluAsnLysLysProGlyGly..... 47

|||||
212 GCAGCCCTGCAGANTGACAGGAAGACGCCCTGTGGGAGGAGGTGAGG 261
47
262 GCCACCAAGCCTGCTGCAACCAACCAACCAACCAACCAACCAACCAAC 311
48AsnValArgArgLysValArgLeuValProAsnPheLeuTr 62
312 AACCAACCAAGAGAGCGGCTGCGGACTGCGCCCTAACTTCCGATG 361
62 pAlaIleProAsnArgHisValAspHisSerGluGly.....GlyGluG 77
|||||
362 GGCATTCCCAACAGGAGATGAATGACGGTGGTGGGAGATGGAGATG 411
77 luValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
|||||
412 ATATGGAAATCTCATGGAGGATGAGAGATCCGGAAGAGCTTAGG 461
94 GluGlnGlnMetArgProTyrThrArgPheArgThrProGluProAspAs 110
|||||
462 GAGCTACAGCTGAGAAATGCTACGCATCCTTATGGGGAGCTGTCTAA 511
110 nHisTyrAsp.....PheCysLeuIlePro 118
|||||
512 CCACCAAGATCACCATGATGAATTCGCCCTATGCCT 548

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AA01004

seq_documentation_block:

ID_ AAC01004 standard; cDNA; 485 BP.

XX AAC01004;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 1002.

KW Humao; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR P-FSDB; AAG00998.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 1; SEQ ID 1002; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

OS Homo sapiens.
XX CN1296969-A.
XX PD 30-MAY-2001.
XX PF 23-NOV-1999; 99CN-0124078.
XX PR 23-NOV-1999; 99CN-0124078.
XX PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
XX PI Mao Y, Xie Y;
XX DR WPI: 2001-483897/53.
XX DR P-PSDB; AAG66407.
XX PT Polypeptide-human X chromosome linked gene expression protein 14 and
XX PT polynucleotide for coding said polypeptide -
XX PS Claim 6; Page 24 (Disclosure); 31pp; Chinese.
XX CC The present sequence is the coding sequence for human X chromosome linked
XX CC gene expression protein 14. The protein and coding sequence are useful
XX CC for treating diseases e.g. cancer and HIV infection.
XX SQ Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;

alignment_scores:
Quality: 111.00 Length: 132
Ratio: 1.461 Gaps: 6
Percent Similarity: 57.576 Percent Identity: 31.818

alignment_block:

US-09-327-750D-34 x AAH75810 ..
Align seg 1/1 to: AAH75810 from: 1 to: 767
1 MetalaSerLysValLysGlnValLeuAspLeuThrVal..... 14
101 ATGGAGTCCCAAGAGGAGGAGCGTTAAACAATCTCATCGTGGAAATGT 150
15GLYLSAspLysLysAsnLysGlyGlyLysAlaSerLysG 29
151 CAACGAGAAATGATGAAAGATGAAAGAGGAGCAAGTTCCTAATAA. 199
29 InSerGluGluGluSerHisLysLeuGluValGluAsnLysLysPro 45
200 ..GGGAGGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGTGCT 247
46 GlyGlyAsnValArgArg...LysValArgArgLeuValProAspPheLe 61
248 AGAGAAACCGTAGGCGGTCCCGTTCGAGGAGCCATCCCTGCAGTATAG 297
61 uTPAlaIleProAsnArgHisValasp.....HisSerG 73
298 ATGGGACATAATCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 347
73 LuGlyGlyGluGluValGlyArgPheValGlyLysValMetGluAlaLys 89
348 AGATATATGAAAGGATGGGAGAGGAGTGCAGACAGCTGATGGAAGCTG 397
90 ArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgThrPr 106
398 AGG.....GAAAGCAGGTTGAGTCTCATAGTCTCGGGCAGTCAGCAC 438
106 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118
439 TGATCCCCCTCACCAGTACCATCAGATGAGTTTGCCTTATGCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI58581

seq_documentation_block:
ID_AAI58581 standard; cDNA; 862 BP.
XX AC AAI58581;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 784.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; seq
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
XX P-PSDB; AAM39425.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 784; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.

Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:
Quality: 111.00 Length: 132
Ratio: 1.461 Gaps: 6
Percent Similarity: 57.576 Percent Identity: 31.818

alignment_block:


```

XX SQ Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

alignment_scores:
  Quality: 103.00 Length: 132
  Ratio: 1.411 Gaps: 6
  Percent Similarity: 55.303 Percent Identity: 30.303

alignment_block:
US-09-327-750D-34 x AAC03880
Align seg 1/1 to: AAC03880 from: 1 to: 662
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
207 ATGGAGTCCAAGAGGACGAGCGTTAAACAATCTCATCTGCGAAATGT 256
15 .....GluLysAspLysLysLysLysGlyGlyLysAlaSerLysG 29
257 CAACGAGAAATGATCAAAAGATGAAAAGGAGCAAGTGTCTAATAAA. 305
29 InSerGluGluGluSerHisHisLeuGluGluValGluAsnLysLysPro 45
306 ..GGGAGCCCTTGGCCCTACCTTTGRATGTYTRGTGAATACGTGTGCT 353
46 GlyGlyAsnValArgArg...LysValArgArgLeuValProAsnPhel 61
354 AGAGGAATCTAGCGGTTCCGCTTAGCGAGCCCTCCTGCAGTATAG 403
61 uTrpAlaIleProAsnArgHisValAsp.....HisSerG 73
404 ATGGGATATGATGATAGGCTTGGAGAACCCACAGGCAAGGATGANAGA 453
73 LuGlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAlaLys 89
454 AGAATATGGAAGAGGATTGGGAGGAGTGAGACAGCTGATGGAAGAGCTG 503
90 ArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgThrPr 106
504 AGG.....GAAAAGCAGTTGAGTCACTAGTCTGCGGGCAGTCAGCAC 544
106 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118
545 TGACCCCTCCACCATGACCATCATGATGATGATGTTTCNNWATGCC 590
seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC03879
seq_documentation_block:
AAC03879 standard; cDNA; 698 BP.
AAC03879;
06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 3877.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI

```

```

XX DR WPI; 2000-500381/45.
XX P-PSDB; AAG03873.
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS Claim 1; SEQ ID 3877; 71pp + CD-ROM; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. An ORF has been identified within the
XX CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX CC derived from 30 different tissues. EST sequences usually correspond
XX CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX CC often obtained for isolating cDNA sequences derived from the 5' ends of
XX CC mRNAs and even in those cases where longer cDNA sequences have been
XX CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX CC mRNAs with intact 5' ends and can therefore be used to obtain full length
XX CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX CC gene therapy and chromosome mapping procedures. They are used to obtain
XX CC upstream regulatory sequences and to design expression and secretion
XX CC vectors.
XX SQ Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:
  Quality: 103.00 Length: 132
  Ratio: 1.411 Gaps: 6
  Percent Similarity: 55.303 Percent Identity: 30.303

alignment_block:
US-09-327-750D-34 x AAC03879
Align seg 1/1 to: AAC03879 from: 1 to: 698
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
243 ATGGAGTCCAAGAGGACGAGCGTTAAACAATCTCATCTGCGAAATGT 292
15 .....GluLysAspLysLysLysLysGlyGlyLysAlaSerLysG 29
293 CAACGAGAAATGATGAAAAGATGAAAAGGAGCAAGTGTGCTAATAAA. 341
29 InSerGluGluGluSerHisHisLeuGluGluValGluAsnLysLysPro 45
342 ..GGGAGCCCTTGGCCCTACCTTTGRATGTYTRGTGAATACGTGTGCT 389
46 GlyGlyAsnValArgArg...LysValArgArgLeuValProAsnPhel 61
390 AGAGGAATCTAGCGGTTCCGCTTAGCGAGCCCTCCTGCAGTATAG 439
61 uTrpAlaIleProAsnArgHisValAsp.....HisSerG 73
440 ATGGGATATGATGATAGGCTTGGAGAACCCACAGGCAAGGATGANAGA 489
73 LuGlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAlaLys 89
490 AGAATATGGAAGAGGATTGGGAGGAGTGAGACAGCTGATGGAAGAGCTG 539
90 ArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgThrPr 106
540 AGG.....GAAAAGCAGTTGAGTCACTAGTCTGCGGGCAGTCAGCAC 580
106 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118
581 TGACCCCTCCACCATGACCATCATGATGATGATGTTTCNNWATGCC 626
seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAI60367
seq_documentation_block:

```

```

Align seg 1/1 to: AA160367 from: 1 to: 858

1  MetAlaSerLysValLysGlnValIleuAspLeuThrVal.....14
   ||| ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
215 ATGGAGTCCAAAGAGAACGAGCGCTTAACCAATCTCATCGTGGAAATGT 264
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
15  ....GluLysAspLysLysAsnLysLysGlyGlyLysAlaSerLysG 29
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
265 CAACCAAGGAAATGATGAAAGAAGATGAAAGGACGAGTTCCTAATAAA. 313
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
29  InSerGluGluSerHisHisLeuGluGluValGluAsnLysLysPro 45
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
314 .GGGAGCGCCTTGCCCTACTCTTGAATGTTAGTAATCTGTGTGCT 361
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
46  GlyGlyAsnValArg.....LysValArgLeuValProAsnPheLe 61
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
362 AGAGGAAACCGTAGCGGTTCCGCGTTAGGCAGCCCATCTGCATATAG 411
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
61  uTPAlaIleProAsnArgHisValasp.....HisSerG 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 ATGGGACATAATGCATAGCTGTGGAGAGCCACAGGCAAGGATGAGAGG 461
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
73  luGlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAlaLys 89
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
462 AGAATATGAAAGGATTGGGAGGAGGCTGAGACAGCTGATGGAAGAGCTG 511
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
90  ArgHisSerLysGluGlnIleMetArgProTyrThrArgPheArgThr 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
512 AGG.....GAAACACGATCTGAGTCATAGTCTCGGGCGACTCAGCAC 552
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
106 oGluPro.....AspAsnHisTyrAspPheCys 115
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
553 TGATCCCCCTCACCATGACCATCAGCATGAGTTTTC 589
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

OM of: US-09-327-750D-34 to: GenEmbl.* out_format : pfs
 Date: Mar 11, 2002 3:34 PM
 About: Results were produced by the GenCore software, version 4.5.
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODE=frame p2n.model -DEV=xlh
 -O=/sgn2/1/USPTO_spool/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOAPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCAL=200 -OUTPMT=pfs -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTPMT=pfs -NORM=ext -HEAPSIZ=500
 -MINLEN=0 -MAXLEN=2000000000 -USER=US09327750_0CGN1_1_8673
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-09-327-750D-34
 Query length: 118
 Database: GenEmbl.*
 Database sequences: 1472140
 Database length: 34134837
 Search time (sec): 4557.230000

score list:	Strd	Orig	ZScore	EScore	Len	Documentation
gb.pr:AK000959	+	372.00	661.44	1.48-28	1229	AK000959 Homo sapiens CDNA FL3
gb.pr:ALX100231	+	372.00	660.55	1.6e-28	1364	ALX100231 Sequence 32 from Pat
gb.pr:HS6353G19	+	372.00	627.19	1.2e-26	6948	AL035494 Human DNA sequence
gb.pr:HUMOGC	+	171.50	302.23	1.5e-08	891	M38188 Human unknown protein f
gb.pr:AF187064	+	171.50	302.23	1.5e-08	891	AF187064 Homo sapiens p75NTR-as
gb.pr:BC003190	+	170.00	300.51	1.8e-08	793	BC003190 Homo sapiens, p75NTR-a
gb.pr:AF187065	+	148.00	264.39	1.9e-06	519	AF187065 Rattus norvegicus p75N
gb.sts:G72708	+	145.00	260.75	3.0e-06	421	G72708 MARC 4953-4954:99193031
gb.ro:AF187066	+	145.00	256.44	5.2e-06	700	AF187066 Mus musculus p75NTR-as
gb.ro:AF097440	+	145.00	254.75	6.5e-06	854	AF097440 Mus musculus brain exp
gb.pr:HSV351F8	+	138.00	208.36	0.0025	45678	AF097440 Mus musculus brain exp
gb.sts:G55294	+	118.50	211.86	0.0016	477	G55294 human STS SHGC-37409, se
gb.ro:AF097439	+	116.50	198.50	0.0044	785	AF097439 Mus musculus brain exp
gb.pr:AF183416	+	112.50	198.50	0.0088	642	AF183416 Homo sapiens ovarian g
gb.pr:AF220189	+	112.50	196.35	0.0117	828	AF220189 Homo sapiens uncharact
gb.pr:HS1988P4	+	112.50	164.41	0.6998	35714	AF008708 Human DNA sequence f
gb.pr:AF251053	+	111.00	194.55	0.0147	744	AF251053 Homo sapiens X-linked
gb.pr:AF207822	+	111.00	192.95	0.0180	898	AX078272 Sequence 76 from Paten
gb.pr:HSV870H8	+	111.00	162.82	0.8586	31321	AF078272 Sequence 76 from Paten
gb.pr:ALX33348	+	111.00	160.62	1.14	40584	ALX33348 Human DNA sequence f
gb.pr:HS080B1	+	111.00	160.53	1.15	41029	AL022169 Homo sapiens chromos
gb.pr:AF237783	+	109.50	191.32	0.0222	791	AF237783 Homo sapiens brain-exp
gb.ro:AF097438	+	105.00	182.74	0.0668	835	AF097438 Mus musculus brain exp
gb.ro:AF051347	+	105.00	182.71	0.0670	838	AF051347 Mus musculus REX-3 mRN
gb.ro:AF097437	+	105.00	174.26	0.1981	2269	AF097437 Mus musculus Bex1 pr
gb.pr:HS714B7	+	102.50	137.78	21.32	98274	I299755 Human DNA sequence fr
gb.pr:AF071813	+	96.00	163.24	0.8138	1225	AF071813 Streptococcus pneumon
gb.pr:AC023561	+	95.50	120.50	195.44	169808	AC023561 Homo sapiens chromo
gb.pr:AC074348	+	95.00	120.24	202.20	157511	AC074348 Homo sapiens clone
gb.pr:AC048344	+	95.00	119.13	232.92	179380	AC048344 Homo sapiens chromo
gb.pr:AC023341	+	94.50	118.72	245.66	169365	AC023341 Homo sapiens chromo
gb.pr:HSAC000373	+	94.00	130.14	56.75	39602	AC000373 Human cosmid g1862d2
gb.pr:HSAC000367	+	94.00	129.38	62.62	43349	AC000367 Human Cosmid g1862d2
gb.pr:AC012232	+	94.00	117.57	284.63	173337	AC012232 Homo sapiens clone
gb.pr:HSV353P9	+	92.50	118.28	259.95	116561	AL023653 Human DNA sequence
gb.pr:ALJ355809	+	92.50	118.01	269.13	120340	ALJ355809 Human DNA sequence
gb.pr:AF195953	+	92.50	113.34	489.70	208618	AF195953 Homo sapiens membra
gb.in:CSA7431	+	92.00	151.74	3.56	2029	AJ007431 Cupienius salei mRNA
gb.pr:AL512842	+	92.00	114.96	397.64	154889	AL512842 Homo sapiens chrom
gb.pr:AC091457	+	92.00	112.30	559.73	212080	AC091457 Mus musculus clone

gb.htg:AL139096 + 91.50 115.47 372.68 131198 I AL139096 Homo sapiens chr
 gb.pr:U72787 + 91.00 131.18 49.69 18510 I U72787 Homo sapiens cosmid
 gb.pr:AL359643 + 91.00 112.53 543.53 166863 I AL359643 Human DNA sequen
 gb.pr:HS678N14 - 91.00 111.78 597.92 182150 I AL109654 Human DNA sequen
 gb.htg:BI011H02 + 90.50 119.60 219.49 65190 I AL442109 Oryza sativa chro

seq_name: gb.pr:AK000959
 seq_documentation_block: 1229 bp mRNA PRI 22-FEB-2000
 LOCUS AK000959
 DEFINITION Homo sapiens CDNA FLJ10097 fls, clone HEMBA1002458, weakly similar
 to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.
 ACCESSION AK000959
 VERSION 1
 KEYWORDS GI:7021945
 SOURCE Oligo mapping: fls (full insert sequence).
 mRNA, clone lib:HEMBAL clone:HEMBAL1002458, 10 weeks whole embryo, mainly head CDNA to
 ORGANISM Homo sapiens

REFERENCE 1 (sites)
 AUTHORS Eukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
 Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
 Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
 Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Oshima,A.
 NEDO human cDNA sequencing project
 JOURNAL Unpublished (2000)

REFERENCE 2 (bases 1 to 1229)
 AUTHORS Isogai,T. and Otsuki,T.
 TITLE Direct Submission
 JOURNAL Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao

Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@helix.co.jp,
 Tel:81-438-52-3951, Fax:81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing and clone selection:
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.

FEATURES Location/Qualifiers
 source 1..1229
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEMBA1002458"
 /clone_lib="HEMBA1"
 /dev_stages="embryo, 10 weeks"
 /tissue_type="whole embryo, mainly head"
 /note="cloning vector: pME18SFL3"
 CDS 191..553
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAA91443.1"
 /db_xref="gi:7021946"
 /translation="MESKEELANLNAGNAQNEGGQAPQNEEERHLGGGEGQ
 KPGGNIRGRVRRVLPNFRWPAIPNRHIEHNEARDVDFVFGQWMEIKRKRREQQMRHY
 MRQPEPDNDHYDCLIP"

BASE COUNT 351 a 243 c 296 g 339 t
 ORIGIN
 alignment_scores:
 Quality: 372.00 Length: 120
 Ratio: 3.758 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:
 US-09-327-750D-34 x AK000959
 Align seg 1/1 to: AK000959 from: 1 to: 1229

details see <http://bacpac.med.buffalo.edu/VECTOR: pcYPAC2>
 IMPORTANT: This sequence is not the entire insert of clone 635G19.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring submissions.
 The true left end of clone U65A4 (281014) is at 69549 in this
 sequence. The true right end of clone U101D3 (285997) is at 100 in
 this sequence.

FEATURES Location/Qualifiers

source

1..69648
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"

/map="q22.1-22.3"
 /clone="RP4-635G19"
 /clone.lib="RCI-4"
 439..483

repeat_region
 /note="TA81 repeat: matches 707..749 of consensus"
 501..561
 repeat_region
 /note="AluJb repeat: matches 137..311 of consensus"
 662..897
 repeat_region
 /note="118 copies 2 mer aa 70% conserved"
 915..1331
 repeat_region
 /note="W5TA repeat: matches 1..419 of consensus"
 1334..1517
 repeat_region
 /note="92 copies 2 mer ga 77% conserved"
 1927..2020
 repeat_region
 /note="MIR repeat: matches 109..217 of consensus"
 2066..2262
 repeat_region
 /note="L1ME2 repeat: matches 5825..6022 of consensus"
 2263..2574
 repeat_region
 /note="AluSp repeat: matches 1..313 of consensus"
 2575..2819
 repeat_region
 /note="L1ME2 repeat: matches 5583..5825 of consensus"
 2821..3039
 repeat_region
 /note="Alu repeat: matches 1..311 of consensus"
 3104..3384
 repeat_region
 /note="L1ME2 repeat: matches 5304..5591 of consensus"
 3391..4007
 repeat_region
 /note="L1MB1 repeat: matches 5558..6162 of consensus"
 4012..4137
 repeat_region
 /note="Tigger2a repeat: matches 11..123 of consensus"
 4141..4252
 repeat_region
 /note="L1 repeat: matches 2920..3031 of consensus"
 4265..4350
 repeat_region
 /note="HY4 repeat: matches 1..87 of consensus"
 4350..4678
 repeat_region
 /note="MER74A repeat: matches 28..369 of consensus"
 4914..5095
 repeat_region
 /note="MER3A repeat: matches 8..189 of consensus"
 5224..5555
 repeat_region
 /note="L2 repeat: matches 2423..2746 of consensus"
 6040..6086
 repeat_region
 /note="L2 repeat: matches 1676..1720 of consensus"
 7413..7478
 repeat_region
 /note="Alu repeat: matches 239..304 of consensus"
 7628..7928
 repeat_region
 /note="AluJb8 repeat: matches 1..316 of consensus"
 7952..7987
 repeat_region
 /note="18 copies 2 mer aa 81% conserved"
 7990..8305
 repeat_region
 /note="AluJb repeat: matches 1..309 of consensus"
 8436..8538
 repeat_region
 /note="MIR repeat: matches 99..212 of consensus"
 10836..11442
 repeat_region
 /note="L1MD repeat: matches 1357..1972 of consensus"
 11443..11571
 repeat_region
 /note="FLAN_A repeat: matches 2..129 of consensus"
 11572..11946
 repeat_region
 /note="L1MD repeat: matches 977..1357 of consensus"
 11947..12245
 repeat_region
 /note="AluX repeat: matches 1..301 of consensus"
 12246..13098

repeat_region
 /note="L1MD repeat: matches 275..977 of consensus"
 13094..13223
 repeat_region
 /note="L1MD repeat: matches 17..113 of consensus"
 13523..13820
 repeat_region
 /note="MLTID repeat: matches 177..505 of consensus"
 13823..13932
 repeat_region
 /note="55 copies 2 mer ct 77% conserved"
 13936..13977
 repeat_region
 /note="MLTID repeat: matches 201..241 of consensus"
 13978..14240
 repeat_region
 /note="AluSg1 repeat: matches 1..270 of consensus"
 14245..14534
 repeat_region
 /note="AluSc repeat: matches 1..290 of consensus"
 14535..14721
 repeat_region
 /note="MLTID repeat: matches 2..201 of consensus"
 14724..15604
 repeat_region
 /note="L1MC/D repeat: matches 4646..5565 of consensus"
 15605..15909
 repeat_region
 /note="AluX repeat: matches 24..312 of consensus"
 16069..16363
 repeat_region
 /note="AluX repeat: matches 1..291 of consensus"
 16482..16641
 repeat_region
 /note="AluJb repeat: matches 136..295 of consensus"
 16671..16953
 repeat_region
 /note="AluX repeat: matches 1..284 of consensus"
 17164..17280
 repeat_region
 /note="L1P repeat: matches 4499..4619 of consensus"
 17326..17377
 repeat_region
 /note="26 copies 2 mer at 79% conserved"
 17704..17863
 repeat_region
 /note="AluSc repeat: matches 118..295 of consensus"
 17864..18175
 repeat_region
 /note="AluJb8 repeat: matches 1..318 of consensus"
 18176..18298
 repeat_region
 /note="AluSc repeat: matches 1..118 of consensus"
 18488..18873
 repeat_region
 /note="13 copies 2 mer tg 100% conserved"
 18876..19171
 repeat_region
 /note="AluSg repeat: matches 1..295 of consensus"
 19183..19232
 repeat_region
 /note="25 copies 2 mer ac 100% conserved"
 19246..19533
 repeat_region
 /note="AluY repeat: matches 5..292 of consensus"
 19539..19837
 repeat_region
 /note="MER74B repeat: matches 61..368 of consensus"
 20016..20194
 repeat_region
 /note="MER74A repeat: matches 295..487 of consensus"
 20394..20692
 repeat_region
 /note="AluSg repeat: matches 1..302 of consensus"
 20701..20809
 repeat_region
 /note="LTR1 repeat: matches 682..785 of consensus"
 20877..21061
 repeat_region
 /note="LTR1 repeat: matches 1..195 of consensus"
 21186..21636
 repeat_region
 /note="LTR29 repeat: matches 3..498 of consensus"
 21609..21737
 repeat_region
 /note="MER34 repeat: matches 409..543 of consensus"
 21771..22076
 repeat_region
 /note="AluY repeat: matches 2..308 of consensus"
 22878..22974
 repeat_region
 /note="MIR repeat: matches 92..191 of consensus"
 23452..23582
 repeat_region
 /note="FLAN_C repeat: matches 1..131 of consensus"
 23583..23624
 repeat_region
 /note="21 copies 2 mer ta 76% conserved"
 23735..24041
 repeat_region
 /note="AluSg1 repeat: matches 1..308 of consensus"
 24045..24072
 repeat_region
 /note="14 copies 2 mer tt 89% conserved"
 25274..25582
 repeat_region
 /note="AluJb repeat: matches 1..289 of consensus"
 25583..25747
 misc_feature
 /note="match: GSS B63622"


```

|||||
162 AACAAATCTCATGCGCAATGTCACACAGGAAACGAAGATGGA 211
34 rHisHisLeuGluValGluAsnLysLysProGlyGly..... 47
212 GCAGCCCTGCAGATGGACAGGAGACCGCCCTGTGGGAGGTGAGG 261
47 .....
262 GCCACGAGCTGCTGCAACAAACAAACAAACAAACAAACAAACAC 311
48 .....AsnValArgArgLysValArgLeuValProAsnPheLeuTr 62
312 AACACACGAGAGCGGCGGCTGCGGACCTGCGCCCTAACTCCGATG 361
62 palatleProAsnArgHisValAspHisSerGlyGly.....GlyGluG 77
362 GGCCATTCACACAGGAGATGATGACGGTTGGTGAGATGAGATG 411
77 luValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
412 ATATGGAAATGTTATGGAGGAGATGAGAGATCCGGAGAAAGCTTAGG 461
94 GluGlnMetArgProTyrThrArgPheArgThrProGluProAspAs 110
462 GAGCTACAGCTGAGAAATGCTACGCATCTTATGGGGAGCTGCTAA 511
110 nHistyTasp.....PheCysLeuLeuPro 118
512 CCACGAGATCACCATGATGATTCGCTTATGCCT 548
seq_name: gb_ro:AF097440

```

```

seq_documentation_block:
LOCUS AF097440 854 bp mRNA ROD 13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 3 (Bex3) mRNA,
complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 854)
Brown, A.L. and Kay, G.F.
Bex3, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
JOURNAL 99172070
MEDLINE
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE 2 (bases 1 to 854)
AUTHORS Brown, A.L. and Kay, G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
SOURCE
1. 854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosome="X"
/map="near Plp"
/tissue_type="pooled organs"
1. 854
/gene="Bex3"
172. 546
/gene="Bex3"
/codon_start=1
/product="brain expressed X-linked protein 3"
/protein_id="AAD24431.1"

```

```

/db_xref="GI:4580594"
/translation="MANVHOENEEPEQPLANGQEDRPVGGEGHQPAANNNNNNHHN
NHHRRGOARRLAPFNFWAIPNROMDGLGDDGDMFEMFEMREIRRLKRLRLRNC
LRLIMGELSNHHHDHDFCLMP"
BASE COUNT 237 a 212 c 228 g 177 t
ORIGIN

alignment_scores:
Quality: 145.00 Length: 129
Ratio: 2.014 Gaps: 3
Percent Similarity: 55.814 Percent Identity: 28.682

alignment_block:
US-09-327-750D-34 x AF097440 ..
Align seg 1/1 to: AF097440 from: 1 to: 854

18 LysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGluSe 34
157 AAACAATCTCATGCGCAATGTCACACAGGAAACGAAGATGGA 206
34 rHisHisLeuGluValGluAsnLysLysProGlyGly..... 47
207 GCAGCCCTGCAGATGGACAGGAGACCGCCCTGTGGGAGGTGAGG 256
47 .....
257 GCCACGAGCTGCTGCAACAAACAAACAAACAAACAAACAAACAC 306
48 .....AsnValArgArgLysValArgLeuValProAsnPheLeuTr 62
307 AACACACGAGAGCGGCGGCTGCGGACCTTGCCTTAACCTCCGATG 356
62 palatleProAsnArgHisValAspHisSerGlyGly.....GlyGluG 77
357 GGCCATTCACACAGGAGATGATGACGGTTGGTGAGATGAGATG 406
77 luValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
407 ATATGGAAATGTTATGGAGGAGATGAGAGATCCGGAGAAAGCTTAGG 456
94 GluGlnMetArgProTyrThrArgPheArgThrProGluProAspAs 110
457 GAGCTACAGCTGAGAAATGCTACGCATCTTATGGGGAGCTGCTAA 506
110 nHistyTasp.....PheCysLeuLeuPro 118
507 CCACGAGATCACCATGATGATTCGCTTATGCCT 543
seq_name: gb_pr:HSV351F8

seq_documentation_block:
LOCUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V351F8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION 270719
VERSION 270719.1 GI:1261915
KEYWORDS X.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 45678)
Whiteley, M.
Direct Submission
Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RO, UK. E-mail enquiries: humquery@sanger.ac.uk
IMPORTANT: This sequence is the entire insert of clone V351F8. The
true left end of clone V351F8 is at 1 in this sequence. The true
right end of clone V351F8 is at 45678.
V351F8 is from the human chromosome X-specific cosmid library.
Location/Qualifiers
FEATURES

```



```

primer_bind      complement(130..149)
BASE COUNT      153 a 115 c 79 g 130 t
ORIGIN

alignment_scores
  Quality: 118.50      Length: 83
  Ratio: 2.370        Gaps: 2
  Percent Similarity: 60.241      Percent Identity: 33.735

alignment_block:
US-09-327-750D-34 x G35294/rev ..
Align seg 1/1 to reverse of: G35294 from: 1 to: 477

43  LyslyaproclyGlyAsnValArgArgLysValArgArgLeuValProAs 59
      :::::::::::::::::::::
474 CAGAGCGCTGGAGGAATGTTAAAGGGGTTGGGCTCCACCTGCCCGGG 425

59  npeLeuTrpAlaIleProAsnArgHisValasphHisSerGlu..... 73
      :::::::::::::::::::::
424 TTTTGGAGAGGAGTGTGCCCAATAGCTTGTGATAACATTGATATGATAG 375

74  ....GlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAla 88
      :::::::::::::::::::::
374 ATGGAGATGGAGATGATATGGAACGGTTCATGGAGGAGATGAGAGAGCTA 325

89  LysArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgTh 105
      :::::::::::::::::::::
324 AGGAGGAAATATAGGAACCTTCAGTTGAGGTACAGCTGCGCATCTTAT 275

105 rProGluPro.....AspAsnHistyrAspPheCysLeuIlePro 118
      :::::::::::::::::::::
274 AGGGACCTCTCCACCATGATCATCATGATGAGTTTGCCTTATGCT 226

seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS      AF097439      785 bp      mRNA      13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 2 (Bex2) mRNA,
            complete cds.
VERSION    AF097439
ACCESSION AF097439.1 GI:4580591
KEYWORDS   house mouse.
SOURCE     Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS    Brown,A.L. and Kay,G.F.
TITLE      Bex1, a gene with increased expression in parthenogenetic embryos,
            is a member of a novel gene family on the mouse X chromosome
JOURNAL    Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE    99172070
REMARK     Erratum:{{published erratum appears in Hum Mol Genet 1999
            May;8(5):943}}
            2 (bases 1 to 785)
REFERENCE   Brown,A.L. and Kay,G.F.
AUTHORS     Direct Submission
TITLE      Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
            Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
            Location/Qualifiers
            1..785
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /db_xref="dbEST:W48832"
            /chromosome="X"
            /map="near Pip"
            /dev_stage="embryo; 15.5 dpc"
            1..785
            /gene="Bex2"
            139..528
            CDS
gene
CDS

```

Tue Mar 12 09:01:42 2002

TITLE Direct Submission
 JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China

FEATURES
 source
 1. 642
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="adrenal gland"
 3. 380
 /note="HGR74-h protein."
 /codon_start=1
 /evidence=not_experimental
 /product="ovarian granulosa cell 13.0 kDa protein hGR74 homolog"
 /protein_id="AA09685.1"
 /db_xref="GI:9963771"
 /translation="MESKEKRAVNSLSENAOENEEKQVANKGEPLALPLDAGEYCVPRGNRRFRVROPIYVYRWDLHVRVGPQGRMRRENVQREGDDVRQLMEKLRE
 HSLRAVSTDPFHDDHDEFLMP"

CDS
 204 a 118 c 157 g 163 t
 BASE COUNT
 ORIGIN

alignment_scores
 Quality: 112.50 Length: 128
 Ratio: 1.424 Gaps: 5
 Percent Similarity: 61.719 Percent Identity: 28.906

alignment_block
 US-09-327-750D-34 x AF183416 ..
 Align seg 1/1 to: AF183416 from: 1 to: 642

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLys... 16
 3 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGCTCAGCATGGAATGC 52
 17 AspLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluG 33
 53 CAACCAAGAAATGAGAAAGAGCAAGTGTCTAATAAGGGGAGCCCT 102
 33 LuSerHisHisLeuGluGluValGluAsnLysLysProGlyLysAsnVal 49
 103 TGGCCCTCCCTTGGTGTGTAATACTGTGCTAGAGAAATCGT 152
 50 ArgArg...LysValArgArgLeuValProAsnPheLeuTIPAlaIlePr 65
 153 AGCGGTTCCTGCTAGCAGCCCTCCTGCTAGTATAGTGGATATGAT 202
 65 oAsnArgHisValAsp.....HisSerGluGlyGluGluG 77
 203 GCATAGGCTTGGAGAACACAGCAAGGATCAGAGAAGAAATATGAAA 252
 77 LuValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
 253 GGATGGGAGAGGTGAGACAGCTGATGGAAGAGCTGAGG..... 293
 94 GluGlnGlnMetArgProTyrThrArgPheArgThrProGluPro..... 108
 294 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCAGCTGACCCCTCA 343
 109AspAsnHisTyrAspPheCysLeuIlePro 118
 344 CCATGACCATCATGATGAGTTTGCCTATGCCCC 377

seq_name: gb_pr:AF220189
 seq_documentation_block:
 LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
 DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA, complete cds.
 ACCESSION AF220189

/gene="Bex2"
 /codon_start=1
 /product="brain expressed x-linked protein 2"
 /protein_id="AA024430.1"
 /db_xref="GI:4580592"
 /translation="MESKVQGVKNLNMENDHOEKEEKEEKPDQASKRDPVALPFEA
 GDYVVPGRGRRFRVROPIYVYRWDLHVRVGPQGRMRRENVQREGDDVRQLMEKLRE
 ROLSHSLRAVSTDPFHDDHDEFLMP"

BASE COUNT 212 a 163 c 235 g 175 t
 ORIGIN

alignment_scores
 Quality: 116.50 Length: 130
 Ratio: 1.553 Gaps: 5
 Percent Similarity: 57.692 Percent Identity: 30.000

alignment_block
 US-09-327-750D-34 x AF097439 ..
 Align seg 1/1 to: AF097439 from: 1 to: 785

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
 139 ATGGAGTCCAAAGTGAACACAGCGGAGAAATCTCAACATGGAGATGA 188
 15GluLysAspLysLysAsnLysLysGlyLysAlaSerLysG 29
 189 CCATCAGAAAGAGGAGAAAGAGAAAGCCACAGGATGCTAGCAAAA 238
 29 InSerGluGluGluSerHisHisLeuGluGluValGluAsnLysLysPro 45
 239 GGGATCGGATGTCGCGCCCTTCGAAGCTGGAGACTACTACGTGCT 288
 46 ...GlyGlyAsnValArgLysValArgArgLeuValProAsnPheLe 61
 289 AGAGGAGTCCGAGCGGCTTCGGGTTCGGCAGCCATGTCACATACAG 338
 61 uTPAlaIleProAsnArgHisValAspHisSerGluGly.....G 75
 339 ATGGGACCTGATCATAGG...GTTGGGAGCCCGGAGGAGATGAGAG 385
 75 LyGluGluValGlyArgPheValGlyGlnValMetGluAlaLysArgHis 91
 386 AGGAGACGTACAGAGGTTGGGGATGATCTGAGACAGCTCATGAGAAG 435
 92 SerLysGluGlnGlnMetArgProTyrThrArgPheArgThrProGluPr 108
 436 CTGAGGGAAGGCGAGCTGAGCCACAGCCCTGCGGGGGTTAGCAGCTGACCC 485
 108 o.....AspAsnHisTyrAspPheCysLeuIlePro 118
 486 GCCTCATCATGACCAACCATCATGAGTTTGCCTATGCCCC 525

seq_name: gb_pr:AF183416
 seq_documentation_block:
 LOCUS AF183416 642 bp mRNA PRI 02-SEP-2000
 DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog mRNA, complete cds.
 ACCESSION AF183416
 VERSION AF183416.1 GI:9963770
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 642)
 Peng, Y., Qian, b., Tu, Y., Xu, S., Han, Z., Fu, G. and Chen, Z.
 A novel gene expressed in human adrenal gland
 Unpublished
 2 (bases 1 to 642)
 Peng, Y., GU, W., Huang, C., Xu, S., Han, Z., Fu, G. and Chen, Z.
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE

OM of: US-09-327-750d-34 to: EST.* out_format: pfs
Date: Mar 11, 2002 2:16 PM
About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frat-p2n model -DEV=xlh
-O=/cgn2_l/US9327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -OPT=fastp -SUFFIX=rat -GAPOF=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.500 -FCAPOP=6.000
-FCAPEXT=0.050 -XGAPOF=10.000 -XGAPEXT=0.500 -DELOP=6.000
-FCAPEXT=7.000 -START=1 -MATRIX=biosum52 -TRANS=human40.cdi
-DEEXT=7.000 -START=1 -MATRIX=biosum52 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0
-ALIGN=5 -MODE=LOCAL -OUTPM=pts -NORM=ext -HEADSIZE=500
-MINLEN=0 -MALEN=2000000000 -USER=US9327750.ecgnl.15854
-NCPU=6 -ICPU=6 -CORELOC -DEV_TIMECUT=120 -WARN_TIMECUT=30
-NO_ALFAP -WARN -THREMS=1

Search information block:

Query: US-09-327-750d-34
Query length: 118
Database: EST.*
Database sequences: 11351937
Database length: 107921985
Search time (sec): 4085.940000

score_list:

seq_id	Strd	Orig	Score	Len	Document
gb_est1:AL527867	568	620.00	1031.44	568	1 A1227867 EST24562 Normalized
gb_est2:BI040889	568	592.00	984.02	568	1 B1289546 UI-R-DKO-cff-c-12-0-UI
gb_est1:AM934883	568	546.00	906.12	568	1 A1010400 Mus musculus ES cells
gb_est2:BG03261	568	546.00	906.12	568	1 BG03261 H308C08-5 NIA Mouse
gb_est1:AA73525	568	534.00	889.85	568	1 AA73525 V978A09.r1 Bierstead M
gb_est1:AM751884	568	520.00	865.45	568	1 AM751884 UI-R-BJO-adn-b-06-0-UI
gb_est2:BG070341	568	510.00	846.61	568	1 BF723075 mab27e12.y1 Soares_NME
gb_est1:BE533774	568	503.00	837.44	568	1 BE533774 us27H06.y1 Soares_NME
gb_est2:BF723075	568	483.00	804.95	568	1 BF723075 mab27e12.y1 Soares_NME
gb_est1:W1712	568	466.00	779.24	568	1 W1712 mab7604.r1 Soares mouse
gb_est1:AM11529	568	456.00	712.42	568	1 AM11529 mab162896.r1 Bedington
gb_est2:W5483	568	456.00	712.42	568	1 W5483 mab162896.r1 Bedington
gb_est2:W5483	568	456.00	712.42	568	1 W5483 mab162896.r1 Bedington
gb_est1:AL1181340	568	410.00	683.49	568	1 AL1181340 v911b236 Bedington m
gb_est2:BF722037	568	407.00	679.16	568	1 BF722037 mab27e12.x1 Soares_NME
gb_est1:AA104150	568	399.00	665.68	568	1 AA104150 m050h03.r1 Life Tech
gb_est2:W83403	568	388.00	647.24	568	1 W83403 m050h03.r1 Life Tech
gb_est1:AA893925	568	386.00	644.71	568	1 AA893925 EST197728 Normalized
gb_est2:BF55502	568	385.00	641.88	568	1 BF55502 UI-R-E1-f1-f-07-0-UI
gb_est1:BG03261	568	373.00	623.91	568	1 BG03261 H308C08-5 NIA Mouse
gb_est2:W83403	568	372.00	619.33	568	1 W83403 m050h03.r1 Life Tech
gb_est1:AL54124	568	372.00	616.23	568	1 AL54124 LTI-NFL006.P1
gb_est1:AU117899	568	372.00	616.08	568	1 AU117899 AU117899 HEMBA1 Homo
gb_est2:BG12197	568	372.00	615.77	568	1 BG12197 603249748F1 NIH_MGC_9
gb_est1:AL54120	568	372.00	615.77	568	1 AL54120 AL54120 LTI-NFL006.P1
gb_est2:BG070427	568	372.00	615.59	568	1 BG070427 602673481F1 NIH_MGC_9
gb_est1:AL560650	568	372.00	615.26	568	1 AL560650 AL560650 LTI-NFL006.P1
gb_est2:BG03261	568	372.00	615.21	568	1 BG03261 H308C08-5 NIA Mouse
gb_est1:AL539004	568	372.00	614.94	568	1 AL539004 AL539004 LTI-NFL006.P1
gb_est2:W83403	568	372.00	614.94	568	1 W83403 m050h03.r1 Life Tech
gb_est1:AL531630	568	367.00	610.61	568	1 AL531630 AL531630 LTI-NFL006.P1
gb_est1:AL523320	568	365.00	603.99	568	1 AL523320 AL523320 LTI-NFL006.P1
gb_est1:AM85267	568	364.00	606.04	568	1 AM85267 RC5-LT0001-261199-011
gb_est2:BG527431	568	364.00	603.94	568	1 BG527431 602552727F1 NIH_MGC_5
gb_est1:BG03261	568	362.00	602.97	568	1 BG03261 H308C08-5 NIA Mouse
gb_est1:AL573904	568	362.00	597.32	568	1 AL573904 AL573904 LTI-NFL006.P1
gb_est2:BI040887	568	360.00	600.82	568	1 BI040887 QV3-NT0278-120201-502

gb_est1:AL582060 - 359.50 594.05 6.3e-24 1037 1 AL582060 AL582060 LTI-NFL
gb_est2:BI040889 + 359.50 593.18 7.1e-24 565 1 BI040889 QV3-NT0278-120201
gb_est1:AM934883 + 355.00 589.55 1.1e-23 668 1 AM934883 RCI-DT0001-230200
gb_est1:AM934905 + 355.00 589.38 1.2e-23 685 1 AM934905 RCI-DT0001-230200
gb_est1:AM934769 + 354.50 586.78 1.2e-23 662 1 AM934769 RCI-DT0001-230200

seq_name: gb_est1:AL227867

seq_documentation_block:
LOCUS A1227867 568 bp mRNA EST
DEFINITION EST24562 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBRCM82 3' end, mRNA sequence.

ACCESSION A1227867

VERSION A1227867

KEYWORDS EST

SOURCE Rattus sp.

ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 568)

AUTHORS Lee, N.H., Glodek, A., Chandrasekhar, I., Mason, T.M., Quackenbush, J.,

Kerlavage, A.R., and Adams, M.D.

TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat

Gene Index

COMMENT Unpublished (1998)

Other ESTs: TC58086

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301) 538-2338

Fax: (301) 538-0208

Email: nhlee@igrr.org

Seq primer: M13-21

Location/Qualifiers

1..568

/organism="Rattus sp."

/db_xref="ATCC (inhost):2036891"

/db_xref="taxon:10118"

/clone="RBRCM82"

/note="Normalized rat brain, Bento Soares"

/note="Organ: brain; Vector: pT73pac; Site: 1; EcoRI;

Site: 2; NotI; 108 g 177 t

BASE COUNT 140 a 143 c 108 g 177 t

ORIGIN

alignment_scores:

Quality: 620.00 Length: 118

Ratio: 5.299 Gaps: 0

Percent Similarity: 99.153 Percent Identity: 99.153

alignment_block:

US-09-327-750d-34 x A1227867/rev ..

Align seg 1/1 to reverse of: A1227867 from: 1 to: 568

1 MetAlaserLysValLysGlnValLysLeuAspLeuThrValGluLysAs 17

|||||

528 ATGGCTCCAAAGTCAACCAAGTCATCTGATCTCCTGCGAGAGAGA 479

|||||

17 pLysLysAsnLysLysGlyLysLysLysLysLysLysLysLysLys 34

|||||

478 CAAGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 429

|||||

34 ATGAG 50

|||||

428 CCCACCATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379

|||||

51 ATGAG 67

|||||

378 AGGAAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 329

|||||

FEATURES
source

Location/Qualifiers
1. 458
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone_image="72056"
/clone_lib="Barstead MPLR81"
/sex="mixed"
/tissue_type="kidney"
/dev_stage="P0.5"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker. Site 1: EcoRI; site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', 3']; double-stranded cDNA was ligated to Eco RI adaptors [CATGATTCGGTACC], digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library constructed by Bob Barstead." 147 a 88 c 115 g 108 t

BASE COUNT
ORIGIN
alignment_scores
Quality: 534.00 Length: 116
Ratio: 4.944 Gaps: 0
Percent Similarity: 93.103 Percent Identity: 86.207

alignment_block:
US-09-327-750D-34 x AA473525

Align seg 1/1 to: AA473525 from: 1 to: 458

3 SerLysValLysGlnValLeuAspLeuThrValGluLysAspLys 19
13 GCGAAATTTAAACAGCTACTGCTCTCAGTGGAGAGACAAAA 62
19 aAnLysValGlyLysAlaSerLysGlnSerGluGluSerHis 36
63 AGACAAAAGGTGGAGAGGCTCCAAACAAAGTGAAGAACCCACC 112
36 lAlaGluLysValGluAsnLysLysProGlyLysValArgLys 52
113 ATCTGGAGAGGTGAACAGAGCTGGGGGAAATGTCGAGGAA 162
53 ValArgLeuValProAsnPheLeuTrpAlaLeuProAsnArgHis 69
163 GTACGCGGACTTGTCTTCTCTGCGCCATACCAATAGGCATGT 212
69 lAspHisSerGluGlyGluGluValGlyArgPheValGlyGlnVal 86
213 TGATCGCATATGAGGGAGAGGATTTGTAGTCAGGGAA 262
86 etGluAlaValArgHisSerLysGluGlnSerArgProThrArg 102
263 CAGAGTCAAGAGAAAGACTACGGAGCAGCGTACAGGCGT 312
103 PheArgThrProGluProAsnHisTyrAspPheCysLeuIlePro 118
313 TTCCGAACCCCGAGACCTGACAATATTACGACTTTTGCCTCATCT 360

seq_name: gb_est1:AW251884

seq_documentation_block:
LOCUS AW251884.1 540 bp mRNA EST 17-DEC-1999
DEFINITION UT-R-B30-adn-b-06-0-UI-3, Rattus norvegicus cDNA clone
UT-R-B30-adn-b-06-0-UI-3, mRNA sequence.
ACCESSION AW251884
VERSION AW251884.1 GI:6595475
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE
1 (bases 1 to 540)
AUTHORS
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
451 Easty O'Connell
151 Research Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoaresblue.wesg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized atrium at 16.5 dpc library cDNA library Preparation:
B. Soares Lab. One distribution: clones will be available
through Research Genetics (www.resgen.com)

Seq primer: M13 Forward
POLYA-Yes.

Location/Qualifiers

1. 540
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone-ut-r-b30-adn-b-06-0-UI-
/clone_lib="UI-R-B30"
/dev_stage="adult"
/adaptor="DH10B (Life Technologies)"
/polylinker="Site 1: Not I, Site 2: Eco RI; The UI-R-B30
library is a subtracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AC1, UI-R-AB1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present at the Not I site
and the oligo-dT track. The library was constructed as
a subtracted library. Contact: Lennan and Soares, Genome Research
6-791-806, 1996.
TAG_Lib=UI-R-B30.
TAG_Tissue=atrium at 16.5 dpc
TAG_SEQ=GATTC

BASE COUNT
ORIGIN
146 a 120 c 94 g 180 t

alignment_scores
Quality: 520.00 Length: 97
Ratio: 5.417 Gaps: 0
Percent Similarity: 98.969 Percent Identity: 98.969

alignment_block:
US-09-327-750D-34 x AW251884/rev

Align seg 1/1 to reverse of: AW251884 from: 1 to: 540

22 LysGlyLysAlaSerLysGlnSerGluGluSerHisLeucl 38
538 AAGGTGGAGAGGCTCCAAACAAAGTGAAGAGATCCCATCTGGA 489
38 uGluValGluAsnLysLysProGlyLysValArgLysValArgA 55
488 AGAGTTTGAACAAAGAGCTGGGGCAATGTCAGGAGGAAGTCAGGC 439
55 rGluValProAsnPheLeuTrpAlaLeuProAsnArgHisValAspHis 71

CELL 13.0 KD PROTEIN HGR74. ; mRNA sequence.
 BE333774
 BE333774.1 GI:9207550

RECORDS	SOURCE	ORGANISM
ES1.	house mouse.	
	Mus musculus	
	Eukaryota; Metazoa;	Chordata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 463)

REFERENCE
1 (bases 1 to 463)
AUTHORS
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

**JOURNAL
COMMENT**

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: crabbs-remail@nih.gov

Email: cgepbs-re@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Ronaldo
Ph.D.,
CDNA Library Arranged by: The I M A C B Consortium (TIMC)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CDAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ATM site.

found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MG1:1063807
Seq primer: -40RP from GIBCO.
Location/Qualifiers
3,150

```
source      ,
1. 463      /organism="Mus musculus"
            /db_xref="taxon:10090"
```

```

/clone="IMAGE:3168347"
/clone_lib="Soares_NMEBA_branchial_arch"
/tissue_type="branchial arches"

```

```

/tissue_type="branchial arches"
/dev_stage="embryo, 10.5 dpc"
/lab_host="DH10B (phage resistant)"
/note="Vector: pT7T3D-pac (Pharmacia) with a modified

```

[illegible]

3'): double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT731 vector. Library

BASE COUNT 155 a 91 c 147 a 70 t

BASE COUNT	155 a	91 c	147 g	70 t
ORIGIN				

alignment_scores:	
Quality:	483.00
PctC:	4.689
Gap:	0
Length:	113

```

Ratio: 4.689      Gaps: 0
Percent Similarity: 91.150      Percent Identity: 84.956

```

alignment_block:
US-09-327-750D:34 x BE333774 ..

Align seg 1/1 to: BE333774 from: 1 to: 463

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17

126 ATGGCATCCCAATTTAAACAAGTCATCTGGATCTCACGTGTGGAGAAAGA 175

17 pLysLYsaAsnLYsLYsGLyGLyLYsAlaSerLYsGLnSerGLuGLu 34
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 176 CAaAAAGAGCAAAaAAAGGTGGAGGCCTCCAACAAaCTGAAGAGAGAAC 225

176 CAAAAAGAGCAGAAAAAGGAGGAGGAGGCCCTCCAAACAGAGAGAGAGAGAGAC 423

34 erHISHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
|||||

236 CCACCATCTCGAAGAGGTTGAAAACAGAGCCTGGGGGAAATGTCCGA 275

51 AatGveVa|AatAatTeuVa|ProAenDheTeuTrnAaT|eProAenAr 67

51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArg 67
|||||
276 AGGAAAGTCAGCGGACTGTGGCCCTAACTTCTCTGGGCCATACCAATAG 325

/clone="IMAGE:538606"
 /clone_lib="Beddington mouse embryonic region"
 /sex="pooled"
 /tissue_type="embryo"
 /dev_stage="7.5dpc"
 /lab_host="DRI29"
 /note="gen: whole embryo; Vector: pCMV-SPORT; Site_1: 5' end; Site_2: 3' end; Cloning strategy: PCR; Oligo dT: Gastrulating embryos were collected at 7.5dpc from C57Bl/6 x DBA matings, excluding embryos that had developed head folds and all extraembryonic tissues. Average insert size: 1.3 kb (range: 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995)."

BASE COUNT 110 a 56 c 92 g 54 t
 ORIGIN

alignment_scores: Quality: 444.00 Length: 100
 Ratio: 4.77 Gaps: 0
 Percent Similarity: 93.000 Percent Identity: 86.000

alignment_block:
 US-09-327-750D-34 x AAL17429

Align seg 1/1 to: AAL17429 from: 1 to: 312

1 MetAserLysValIleValIleLeuAspLeuThrValGluLysAs 17
 |||||
 12 ATGGCATCCAAATTATACAGTCAATGATGATCTCTCTGGGAGAGA 61
 17 PylsValAsnLysGlyGlyLysAlaSerLysGlnSerGluGluGlu 34
 |||||
 62 CAAAGAGACAAAGAGTGGGAGGCTCCAGCAAGAGAGGAGAGAC 111
 34 erHisHisLeuGluGluValGluValGluValLysProGlyGlyAsnValArg 50
 |||||
 112 CCACCATCTGGAGAGGTTGAAACAGAGAGGCTGGGGGAAATGTCCGA 161
 51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
 |||||
 162 AGGAAGTCAGGCAGCTGTGGCTAACTTCTCTGGGCCATACCTAATAG 211
 67 gHisValAspHisSerGluGlyGlyGluValGluValArgPheValGly 84
 |||||
 212 GCATGTTGATCCATCAGAGGAGAGGAGTGTGGGAGATTTGTAGTGC 261
 84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
 |||||
 262 AGGACAGAGTCAAGAGAGAGTACGAGGAGCAGCGTGGGCTTAC 311

seq_name: gb_est1:A1012535

seq_documentation_block: 444 bp mRNA EST 15-JUN-1998
 LOCUS A1012535
 DEFINITION EST206986 Normalized rat placenta, Bento Soares Rattus sp. cDNA
 clone IMAGE:538606, 3' end, mRNA sequence.

ACCESSION A1012535
 VERSION A1012535.1 GI:3226367
 KEYWORDS EST
 SOURCE Rattus sp.
 ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 444)
 AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
 Keriavage,A.R. and Adams,M.D.
 TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat
 Genomic Map
 JOURNAL Unpublished (1998)
 COMMENT Contact: Lee, NH
 The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@tigr.org
 Seq primer: M13-21.

FEATURES

source
 1..444mm="Rattus sp."
 /db_xref="taxon:10118"
 /clone="RPLA54"
 /clone_lib="Normalized rat placenta, Bento Soares"
 /note="Organ: placenta; Vector: pT73pac; Site_1: EcoRI;
 Site_2: NotI"

BASE COUNT 135 a 104 c 81 g 124 t
 ORIGIN

alignment_scores: Quality: 426.00 Length: 80
 Ratio: 5.462 Gaps: 0
 Percent Similarity: 97.500 Percent Identity: 97.500

alignment_block:
 US-09-327-750D-34 x A1012535/rev

Align seg 1/1 to reverse of: A1012535 from: 1 to: 444

39 GluValGluAsnLysPheGlyGlyAsnValArgLysValArgAr 55
 |||||
 444 GAGGTGAAACATGAAGCCTGGGGCAATGTTCAGGAGGAAGTCAGGCG 395
 55 GluValProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 72
 |||||
 394 ACTTGCTTACTTCTCTGGGCCCTACTAATAGCATGTTCATCA 345
 72 erGluGlyGluGluValGluValArgPheValGlyGluValMetGluAla 88
 |||||
 344 GTGAAGGGGAGAGAGGTTGGGAGATTCTAGGCGCAGGTGATGGAAGCC 295
 89 LysArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgTh 105
 |||||
 294 AAGAGAAATCTTAAGGAGCAACAGATGAGGCGCTTACAGCGGTTCCGAA 245
 105 rProGluProAsnHisTyrAspPheCysLeuIlePro 118
 |||||
 244 CCTGACCTGACATCATCATTACGCTTTGGCTCTACTCT 205

seq_name: gb_est2:M35893

seq_documentation_block: 472 bp mRNA EST 14-MAY-1996
 LOCUS M35893
 DEFINITION mc5909.r1 Soares mouse embryo NM013.5 14.5 Mus musculus cDNA
 clone IMAGE:352288.5, similar to SW:RC74_HUMAN Q00994 OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HGR74. [1] ; mRNA sequence.

ACCESSION M35893
 VERSION M35893.1 GI:1317733
 KEYWORDS EST
 SOURCE Mouse mouse
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 472)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-RHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

```
34 erHlaLeuGluGluValGluAenLysLysPcGlyGlyAenValArg 50
|||||
268 CCCACCATCTGGAAGAGGTGAAACAGAGAGCCCTGGGGGAAATGTCGGA 317
|||||
51 ArgLysValArgArgLeuValProAsnPhelLeuTrpAlaIleProAsnAr 67
|||||
318 AGAAGTC-AGCCACTGTGCTTAATCTTCTTGGGCCATACCAATAG 366
|||||
67 gHlaValAspHisSerGlyGlyGluGluValGlyArgPheValGlyG 84
|||||
367 GCATTTGATTCGATGAGGAGGAGAGAGTGTGGAGATTGTGTGC 416
|||||
84 lnValMetGluAlaLysArgHisSerGlyGluGluMetArgPro 99
|||||
417 AGGAACAGAGTCAAGAGAGTACCGAGCAGCAGGTGAGGCT 463
|||||
seq_name: gb_est1:AL118340

seq_documentation_block:
LOCUS AL118340 422 bp mRNA EST 23-SEP-1999
DEFINITION V9112026 Beddington mouse dissected endoderm Mus musculus cDNA
ACCESSION clone_328_13N20 5', mRNA sequence.
VERSION 1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Metazoa; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and
Beddington,R.S.
TITLE Isolation of novel tissue-specific genes from cDNA libraries
representing the individual tissue constituents of the gastrulating
mouse embryo.
JOURNAL Development 121 (8), 2479-2489 (1995)
MEDLINE 95401865
COMMENT
Contact: Ruiz,P., Lehrach,H. and Avner,P.
EBC Mouse Transcript Mapping Consortium
Genoscope - CNS
2, rue Gaston Cremieux, CP 5706, 91057 Evry CEDEX, France
Email: pavner@pasteur.fr
Clone available from Ressourcenzentrum, Heubnerweg 6, D-14059
Berlin, Germany. Web site http://www.rzp.de
Seq primer: CCGGCGGAGATCCCGGT.
Location/Qualifiers
location="Mus musculus"
organism="Mus musculus"
strain="C57Bl6 x DBA"
db_xref="taxon:10090"
clone="528_13N20"
/tissue_type="dissected endoderm"
/dev_stage="7.5 dpc"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI; Cloned
unidirectionally. - High quality sequence only submitted. -
Average insert size: 1.2 kb (range: 0.2 - 2 kb)."
BASE COUNT 128 a 86 c 143 g 65 t
ORIGIN

alignment_scores:
Quality: 407.00 Length: 90
Ratio: 4.788 Gaps: 0
Percent Similarity: 94.444 Percent Identity: 87.778

alignment_block:
US-09-327-750d-34 x AL118340 ..
Align seg 1/1 to: AL118340 from: 1 to: 422
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
```

```
|||||
146 ATGGGATCCCAATTTAAACAGTCAATCTGATCTCTCTGTGGAGAGAG 195
|||||
17 pLysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlu 34
|||||
196 CAAAAAGACAAAAGAGAGGTGGGAGGCTCCAAAACAAAGTGAAGAGAAC 245
|||||
34 erHlaLeuGluGluValGluAenLysLysPcGlyGlyAenValArg 50
|||||
246 CCCACCATCTGGAAGAGGTGAAACAGAGAGCCCTGGGGGAAATGTCGGA 295
|||||
51 ArgLysValArgArgLeuValProAsnPhelLeuTrpAlaIleProAsnAr 67
|||||
296 AGGAATCTCAGCGAGCTTGTGCTTACTTTCTTGGGCCATACCAATAG 345
|||||
67 gHlaValAspHisSerGlyGlyGluGluValGlyArgPheValGlyG 84
|||||
346 GCATTTGATTCGATGAGGAGGAGAGTGTGGAGATTGTGTGTGC 395
|||||
84 lnValMetGluAlaLysArg 90
|||||
396 AGGAACAGAGTCAAGAGA 415
```

```
508 AGCGTTCCGACCCGGAACCTGACATCATACGACTTTGCTCAT 557
117 epro 118
558 ACCT 543

seq_name: gb_est2:BG083261
seq_documentation_block: 792 bp mRNA EST 26-JAN-2001
LOCUS BG083261
DEFINITION H3086C08-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION BG083261
VERSION BG083261.1 GI:12565829
KEYWORDS EST.
ORGANISM Mus musculus.
REFERENCE Kargul G.J., Dudekula D.B., Qian Y., Lim M.K., Jaradat S.A., Tanaka T.S., Carter M.G. and Ko M.S.H.
AUTHORS Verification and Initial Annotation of NIA mouse 15K cDNA clone set
TITLE Unpublished (2001)
JOURNAL Other ESTs: H3086C08-3
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute of Health
3331 Leidos Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@leidosun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://ligun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3086 row: C column: 08
Seq primer: -21M13 Reverse
High quality sequence stop: 792
POLYA-No.

FEATURES
Location/Qualifiers
1..792
/organism="Mus musculus"
/strain="C57BL/6J"
/ab="cDNA"
/clone="H3086C08"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/Note="Vector: pSPORT1. Site 1: SalI; site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of 812.5 embryos, extraembryonic part of 87.5 embryos and 812.5 embryos, extraembryonic part of 87.5 embryos) ovary cDNA libraries. Average size is 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1707-1718;
NOT 159 c 247 a 153 c 210 g 179 t 3 others

BASE COUNT 247 a 153 c 210 g 179 t
ORIGIN

alignment_scores:
Quality: 546.00 Length: 118
Ratio: 4.964 Gaps: 0
```

```
Percent Similarity: 93.220 Percent Identity: 87.288
alignment_block:
US-09-327-750d-34 x BG083261
Align seg 1/1 to: BG083261 from: 1 to: 792

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
|||||
190 ATGGCATCCCAATTAACAAAGTCATCTGATCTCACTGCTGGAGAAGA 239
|||||
17 PLYsLysAsnLYsGlyGlyGlyAlaSerLysGlnSerGluGluGus 34
|||||
240 CAAAAAGACAAAAGAGGGGGGAGGCTCCAAACAAAGTGAAGAGAAC 289
|||||
34 eRfHsHsLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
|||||
290 CCCACATCTGGAGAGAGTGGAAACAGAGCTGGGGGAAATGTCGA 339
|||||
51 ArgLYsValArgGlyGlyValProAsnPheLeuArgPAlaIleProAsnAr 67
|||||
340 AGGAAAGTCAGGCGACTTCTGCTCACTTTCTCTGGGCCATACCAATAG 389
|||||
67 gHsValAspHisSerGluGlyGlyGlyGluValGlyArgPheValGlyG 84
|||||
390 GCATGTGTATCCCAATGAGAGGGGAGAGGATTTGGGAGATTGTAGTGC 439
|||||
84 LnvAlMetGluLysArgHisSerLysGluGluGlnMetArgProTyr 100
|||||
440 AGGGAACAGAGAGTCCAGAGAGAACTACGAGACAGCCAGTGGGCTTAC 489
|||||
101 ThrArgPheArgThrProGluProAsnHisTyrAspPheCysLeuIl 117
|||||
490 AGGCGTTTCCGAACCCGGAACCTGACATCATACGACTTTTGCTCAT 539
|||||
117 epro 118
540 ACCT 543

seq_name: gb_est1:AA473525
seq_documentation_block: 458 bp mRNA EST 18-JUN-1997
LOCUS AA473525
DEFINITION v978a09.t1 Barstead MPRB1 Mus musculus cDNA clone IMAGE:197406 5' similar to SM:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74. ? mRNA sequence.
ACCESSION AA473525
VERSION AA473525.1 GI:2201752
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Zukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Martini, L., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1801
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:511536
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 417.
```

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus 1 to 540)
Bonald M.F. (Lennon, G. and Soares M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
Department of Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 9250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
Normalized atrium at 16.5 dpc library. cDNA Library preparation:
Through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers
1. 540
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0-actn-b-06-0-Ui"
/clone_id="UI-R-BJ0"
/lab_hosts="PHI08 (Life Technologies)"
/notes="vector: pT730-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ0
library is a subtracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AFL, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 15 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and the heart string. The heart string is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG_L1B=UI-R-BJ0
TAG_TISSUE=atrium at 16.5 dpc
TAG_SEQ=GATTC 94 g 180 t
..* a 120 c
BASE COUNT
ORIGIN

alignment_scores:
Quality: 520.00 Length: 97
Ratio: 5.417 Gaps: 0
Percent Similarity: 98.96 Percent Identity: 98.969

alignment block:
US-09-327-730D-34 x AW251884/rev ..
Align seq 1/1 to reverse of: AW251884 from: 1 to: 540

22 LysGlyGlyLysAlaSerLysGlnSerGluGluGluSerHisHisLeuG1 38
538 AAAGGTGGGAAGGCGCTCCAAACAAGTAGAGGAAGATCCACCATCTCGA 489
38 GGUUAGUAGUAnLysLysAprGlyGlyAnValArgArgLysValArg 55
488 KAGGTGTGAACAGAACGCTGGGGGGGATGTCAGAGGAAGATCGAC 439
55 GGUUValProAnSerPheLeuThrPheLeuProAnSerHisValLeuPhe 71

```


(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 158 a 109 c 170 g 81 t 1 others
ORIGIN

alignment_scores:
Quality: 568.50 Length: 113
Ratio: 5.168 Gaps: 1
Percent Similarity: 97.345 Percent Identity: 96.460

alignment_block:

US-09-327-750D-35 x BF723075

Align seg 1/1 to: BF723075 from: 1 to: 519

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
181 ATGGCATCCAAATTTAAACAAGTCATCTGATCTGCTGCGAAGA 230
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
231 CAAAAAGACAAAAAGGTGGGAGCCCTCAACACAAAGTGAAGAAGAC 280
33 roHisLysLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
281 CCCACCATCTGGAAGAGGTGAAACAAGAGCCCTGGGGAAATGTCGGA 330
50 ArgLysValArgArgLeuValProAsnPheLeuTTPAlaIleProAsnAr 66
331 AGGAAGTCAGGCACTTGTCCTACTTCTCTGGCCATACCTAATAG 380
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
381 GCATGTTGATCGCAATGAAGGGGAGAGGATGTTGGGAGATTGTAGTGC 430
83 InGlyThrGluValLysArgLysThrGluGlnGlnValArgProTyr 99
431 AGGAACAGANGTCAGAGAAAGACATCCGAGCAGCAGGTGAGGCTTAC 480
100 ArgArgPheArgThrProGluProAspAsnHisTyrAsp 112
481 AGGCGTTTCCAGCGCGGACCTGACAATCATTTACGAC 519

seq_name: gb_est1:BE333774

seq_documentation_block:
LOCUS BE333774 463 bp mRNA EST 14-JUL-2000
DEFINITION us27h06.y1 Soares_NMEBA_branchial_arch Mus musculus cDNA clone IMAGE:3168347 5' similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74. ; mRNA sequence.

ACCESSION BE333774
VERSION BE333774.1 GI:9207550
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 463)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:

image.llnl.gov/image/html/iresources.shtml

MGI:1063807
Seq primer: -40RP from Gibco.

FEATURES
source

1. 463
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGE:3168347"
/tissue_type="branchial_arch"
/dev_stage="embryo, 10.5 dpc"
/lab_host="DH10B (phage resistant)"
/note="Vector: pT73D-pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCATGATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 155 a 91 c 147 g 70 t
ORIGIN

alignment_scores:
Quality: 544.50 Length: 113
Ratio: 4.950 Gaps: 1
Percent Similarity: 97.345 Percent Identity: 95.575

alignment_block:

US-09-327-750D-35 x BE333774

Align seg 1/1 to: BE333774 from: 1 to: 463

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
126 ATGGCATCCAAATTTAAACAAGTCATCTGATCTGCTGCGAAGA 175
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
176 CAAAAAGACAAAAAGGTGGGAGCCCTCAACACAAAGTGAAGAAGAC 225
33 roHisLysLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
226 CCCACCATCTGGAAGAGGTGAAACAAGAGCCCTGGGGAAATGTCGGA 275
50 ArgLysValArgArgLeuValProAsnPheLeuTTPAlaIleProAsnAr 66
276 AGGAAGTCAGGCACTTGTCCTACTTCTCTGGCCATACCTAATAG 325
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
326 GCATGTTGATCGCAATGAAGGGGAGAGATGTTGGGAGATTGGA GTCC 374
83 InGlyThrGluValLysArgLysThrGluGlnGlnValArgProTyr 99
375 AGGGAACAGANGTCAGAGAAAGACATCCGAGCAGCAGGTGAGGCTTAC 424
100 ArgArgPheArgThrProGluProAspAsnHisTyrAsp 112
425 AAGCGTTTCCAGCGCGGACCTGACAATCATTTACGAC 463

seq_name: gb_est2:W17712

seq_documentation_block:

LOCUS W17712 317 bp mRNA EST 10-SEP-1996
DEFINITION M077604.r1 Soares mouse p3NM19.5 Mus musculus cDNA clone IMAGE:335454 5' similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74. [1] ; mRNA sequence.
ACCESSION W17712
VERSION W17712.1 GI:1292113

KEYWORDS SOURCE

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 317)
Marra M. Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theisinger, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
WGI:216854

TITLE JOURNAL COMMENT

seq_name: gb_est1:AI227867
seq_documentation_block: 568 bp mRNA EST 20-JAN-1999
LOCUS AI227867
DEFINITION EST224562 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBCM482 3' end, mRNA sequence.
ACCESSION AI227867
VERSION AI227867.1 GI:3811754
KEYWORDS EST
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

FEATURES source

Seq primer: mob.REGA+ET
High quality sequence stop: 291.
Location/Qualifiers
1..317
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:335454"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: p7T3D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCGAGTGGAGCGCGCATTTTCTTTTCTTTT 3'].
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
BASE COUNT 108 a 63 c 87 g 59 t
ORIGIN

alignment_scores: Quality: 541.00 Length: 101 Ratio: 5.410 Gaps: 0 Percent Similarity: 99.010 Percent Identity: 98.020

alignment_block:
US-09-327-750d-35 x W17712 ..
Align seg 1/1 to: W17712 from: 1 to: 317

17

LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluPr 33
|||||
1 AAAAAACAAAAAGGTGGGAGGCTCCAAACAAAGTGAAGAGAC 50
|||||

33

OHSHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArgA 50
|||||

51

CCACCATCTGGAAGAGGTGAAACAAAGAGGCTGGGGAAATGTCCGAA 100
|||||

50

qLysValArgLeuValProAsnPheLeuTirpAlaIleProAsnArg 66
|||||

101

CGAAAGTCAGCGACTTGTGCCTAACTTCTCTGGGCCATACCAATAGG 150
|||||

67

HisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValIgl 83
|||||

151

CATGTTGATCGCAATGAAGGGGAGAGGATGTTGGGAGATTCTAGTGCA 200
|||||

83 nGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrA 100
|||||
201 GGAACAGAGTCAAGAGAAAGACTACGAGCAGCAGGTGAGGCTTACA 250
|||||
100 qGArgPheArgThrProGluProAsnHisTyrAspPheCysLeuIle 116
|||||
251 GCGTTTCGAACCCCGAGAGTGCACATCATTAGACATTTTGCCCTATA 300
|||||
117 Pro 117
|||
301 CCT 303

seq_name: gb_est1:AI227867

seq_documentation_block: 568 bp mRNA EST 20-JAN-1999

LOCUS AI227867

DEFINITION EST224562 Normalized rat brain, Bento Soares Rattus sp. cDNA clone

RBCM482 3' end, mRNA sequence.

ACCESSION AI227867

VERSION AI227867.1 GI:3811754

KEYWORDS EST

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 568)

AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,

Kerlavage, A.R. and Adams, M.D.

Rat genome Project: Generation of a Rat EST (REST) Catalog & Rat

Gene Index

Unpublished (1998)

Other ESTs: TC58086

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

Seq primer: M13-21.

Location/Qualifiers

1..568

/organism="Rattus sp."

/db_xref="ATCC (inhost):2036891"

/db_xref="taxon:10118"

/clone="RBCM482"

/clone_lib="Normalized rat brain, Bento Soares"

/note="Organ: brain; Vector: p7T3pac; Site_1: ECORI;

Site_2: NotI"

BASE COUNT 140 a 143 c 108 g 177 t

ORIGIN

alignment_scores:

Quality: 539.50 Length: 118

Ratio: 4.860 Gaps: 1

Percent Similarity: 94.068 Percent Identity: 87.288

alignment_block:

US-09-327-750d-35 x AI227867/rev ..

Align seg 1/1 to reverse of: AI227867 from: 1 to: 568

1 MetalaserLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||

528 ATGGCGTCCAAAGTCAACAAAGTCTACTGATCTCTACTGCGAGAAAGA 479
|||||

16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||

478 CAAGAAAAACAAAAAGGTGGGAGGCTCCAAACAAAGTGAAGAGAAAT 429
|||||

33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||

```

|||||
428 CCCACCTCTGGAGAGGTTGAAACACAGAGGCTGGGGCAATGTCAGG 379
50 ArgLysValArgArgLeuValProAsnPhelTrpAlaIleProAsnAr 66
|||||
378 AGGAAGTCAGGCGCACTGTGCCTAACTTTCTATGGGCCATACCTAATAG 329
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
328 GCATGTTGATCACAGTCAGAGGGGAGAGAGGTTGGAGATTCGTAGGCC 279
83 lnglyThrGluVallyArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
278 AGGTGATGGAGCAAGCAAGAAAGCTCTAAGGAGCAACAGATGAGGCGCTTAC 229
100 ArgArgPheArgThrProGluProAsnHisTyrAspPheCysLeuIl 116
|||||
228 ACGGTTTCCCAACCCCTGAACCTGACATCATTAGCACTTTTGCCCTCAT 179
116 ePro 117
|||||
178 ACCT 175

```

seq_name: gb_est2:BI289546

```

seq_documentation_block: 643 bp mRNA EST 19-JUL-2001
LOCUS BI289546
DEFINITION UI-R-DKO-c12-0-UI.s1 UI-R-DKO Rattus norvegicus cDNA clone
UI-R-DKO-c12-0-UI 3', mRNA sequence.
ACCESSION BI289546
VERSION BI289546
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

```

REFERENCE 1 (bases 1 to 643)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu

```

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized rat kidney pool library cDNA Library Preparation: M.B. Soares Lab Clonics distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1-51,

```

>Ar_rich/Low_complexity
Seq primer: M13 Forward
POLYA=Yes.

```

FEATURES
source

```

1. .643 Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/gb_xref="taxon:10116"
/clone="UI-R-DKO-c12-0-UI"
/clone_lib="UI-R-DKO"
/dev_stage="ADULT"
/lab_host="PH10B (Life Technologies)"
/notes="Vector: pT730-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DKO

```

library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRHP (20%), kidney-nRKP (20%), aorta-nRAP (20%), and placenta-nRPP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, embryonic day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the DKO subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (Ct0s), heart (CS0s), kidney (CU0s), aorta (CW0s), and placenta (CX0s). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CT0), heart (CS0), kidney (CU0), aorta (CW0), and placenta (CX0). The resulting pool of about 10,000 clones represented about 66.6% of the final driver population.

```

TAG LIB-UI-R-DKO
TAG_TISSUE=rat kidney pool
TAG_SEQ=CAAGACTGTC"
BASE COUNT 160 a 149 c 113 g 221 t
ORIGIN

```

```

alignment_scores:
Quality: 511.50 Length: 118
Ratio: 4.608 Gaps: 1
Percent Similarity: 94.068 Percent Identity: 86.441

```

alignment_block:
US-09-327-750D-35 x BI289546/rev ..

```

Align seg 1/1 to reverse of: BI289546 from: 1 to: 643
1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
601 ATGGCGCTCCAAAGTCAACAGTCATATCTGATCTCAGTGGAGAGA 552
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGp 33
|||||
551 CAAGAAAAACAAAAAGGTGGGAG. GCCTCCAAACAAAGTGAAGAAGAT 503
33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
502 CCCACCTCTGGAAGAGGTTGAAAACAAAGAGCTGGGGCAATGTCAGG 453
50 ArgLysValArgArgLeuValProAsnPhelTrpAlaIleProAsnAr 66
|||||
452 AGGAAAGTCAGGCGACTTGTGCCTAACCTTCTATGGGCCATACCTAATAG 403
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
402 GCATGTTGATCACAGTCAGTGAAGGGGAGAGAGGTTGGGAGATTCGTAGGCC 353
83 lnglyThrGluVallyArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
352 AGGTGATGGAAGCAAGAAAGTCTAAGGAGGAGCAACAGATCAGGCGCTTAC 303

```

100 ArgArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeu1116
 302 ACAGGTTTCGGAACCCCTGAACCTGACAAATCATTACGACCTTTGGCTCAT 253

116 epro 117

252 ACCT 249

seq_name: gb_est1:A1117429

seq_documentation_block:

LOCUS A1117429 312 bp mRNA EST 15-NOV-1996
 DEFINITION mouse embryonic region Mus musculus cDNA
 clone IMAGE:538606 5' similar to SW:HG74_HUMAN Q00994 OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HGR74. ; mRNA sequence.

ACCESSION A1117429

VERSION A1117429.1 GI:1672442

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:325342

Seq primer: -40ml3 ET

High quality sequence stop: 287.

Location/Qualifiers

FEATURES

source

1. 312

/organism="Mus musculus"

/strain="C57BL6 x DBA"

/db_xref="taxon:10090"

/clone_image="538606"

/clone_lib="Beddington mouse embryonic region"

/sex="pooled"

/tissue_type="embryo"

/dev_stage="7.5dpc"

/lab_host="DH12S"

/note="Organ: whole embryo; Vector: pCMV-SPORT; Site: 1;

SalI; Site: 2; NotI; Cloned unidirectionally. Primer:

Oligo dr. Gastrulating embryos were collected at 7.5dpc

from C57BL6 x DBA matings, excluding embryos that had

developed head folds and all extraembryonic tissues.

Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).

Referenced in Development 121, 2479-2489 (1995)"

BASE COUNT 110 a 56 c 92 g 54 t

ORIGIN

alignment_scores:

Quality: 503.50 Length: 100

Ratio: 5.086 Gaps: 1

Percent Similarity: 99.000 Percent Identity: 98.000

alignment_block:

US-09-327-750D-35 x A1117429

Align seg 1/1 to: A1117429 from: 1 to: 312

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
 12 ATGGCATCCAAATTTAAACAAGTCATCTGATCTGCTGAGGAAGA 61
 16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGlu 33
 62 CAAAAAGAGACAAAAGGGTGGGAAGGCTCCAAACAAAGTGAAGA 111
 33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49
 112 CCCACCATCTGGAAGAGGTTGAAACAAGAACAGCCCTGGGGAA 161
 50 ArgLysValArgGluValProAsnPheLeuThrPalalleProAsn 66
 162 AGAAAGTCAAGGACGCTTGTGCCTAACTTCTCTGGCCATACCTA 211
 66 gHisValAspArgAsnGluGlyGlyLysValGlyArgPheValVal 83
 212 GCATGTGATCCCATGAGGGGAGAGGATGTTGGGAGATTTGTAG 261
 83 lngLysThrGluValLysArgLysThrGluGlnGlnValArgPro 99
 262 AGGACACAGAGTCAAGAGAAAGACTACGAGCAGCAGAGGTGAG 311
 seq_name: gb_est2:W35893

seq_documentation_block:

LOCUS W35893 472 bp mRNA EST 14-MAY-1996
 DEFINITION mc53g09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:352288 5' similar to SW:HG74_HUMAN Q00994 OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HGR74. [1] ; mRNA sequence.

ACCESSION W35893

VERSION W35893.1 GI:1317733

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 472)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:224088

Seq primer: ETPrimer

High quality sequence stop: 441.

Location/Qualifiers

1. 472

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone_image="352288"

/clone_lib="Soares mouse embryo NbME13.5 14.5"

/sex="unknown"

/tissue_type="embryo"

/dev_stage="13.5-14.5dpc total fetus"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTCGGCGGCCGGAATAATTTTTTTTTTTTTTTTTTTTTT

149 a BASE COUNT

3], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos (total RNA provided by Minoru KO, Wayne
State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo.

M.Fatima Bonaldo. " 71 t

149 a BASE COUNT

found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/lresources.shtml

MGI:1471479
Trace considered overall poor quality
High quality sequence stop: 1.

BASE COUNT	58 a	84 c	58 g	86 t
ORIGIN				

alignment_scores:		
Quality:	479.00	Length: 95
Ratio:	5.207	Gaps: 0
Percent Similarity:	96.842	Percent Identity: 93.684

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 465)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:229324

Seq primer: ExPrimer
High quality sequence stop: 328.

FEATURES

source
1. .465
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:367892"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TCGTACCAATCTGAGTGGAGCGCGCGGAATTTTCTTTTCTTTT
1 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
p7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 147 a 90 c 156 g 72 t

alignment_scores:

Quality: 469.50 Length: 99
Ratio: 4.840 Gaps: 1
Percent Similarity: 97.980 Percent Identity: 96.970

alignment_block:

US-09-327-750D-35 x W54487

Align seg 1/1 to: W54487 from: 1 to: 465

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
168 ATGGCATCCAAATTTAAACAAGTCATCTGATCTCCTGCTGGAGAAAGA 217
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||
218 CAAAAAGACAAAAAGAGGTGGGAAGGCCCTCCAAACAAAGTGAAGAAGAC 267
|||||
33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
268 CCCACCATCTGGAGAGGTGTAACACAAAGCCCTGGGGGAATGTCCGA 317
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTyrPalaIleProAsnAr 66
|||||

318 AGAAAGTC..AGCGACTTGTGCTAACTTTCTCTGGGCCATACCAATAG 366
66 gHisValAspArgAsnGluGlyGlyValGluAspValGlyArgPheValG 83
|||||
367 GCATGTGATCCATCAATGAAGGGGAGAGGATGTTGGGAGATTGTAGTGC 416
|||||
83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgPro 98
|||||
417 AGGAACAGCAAGTCAAGAGAAAGACTAGCGAGCAGCAGGTGAGGCCT 463

seq_name: gb_est1:AL118340

seq_documentation_block:

LOCUS AL118340 422 bp mRNA EST 23-SEP-1999
DEFINITION v9113226 Beddington mouse dissected endoderm Mus musculus CDNA
clone 528_13N20 5', mRNA sequence.

ACCESSION AL118340

VERSION AL118340.1 GI:5920179

KEYWORDS EST_1

SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 422)

AUTHORS Harrison, S.M., Dunwoodie, S.L., Arkell, R.M., Lehrach, H. and
Beddington, R.S.

TITLE Isolation of novel tissue-specific genes from cDNA libraries
representing the individual tissue constituents of the gastrulating
mouse embryo

JOURNAL Development 121 (8), 2479-2489 (1995)

MEDLINE 95401865

COMMENT Contact: Ruiz, P., Lehrach, H. and Avner, P.

ESC Mouse Transcript Mapping Consortium

Genoscope - CNS

2, rue Gaston Cremieux, CP 5706, 91057 Evry CEDEX, France

Email: pavner@pasteur.fr

clone available from Ressourcenzentrum, Heubnerweg 6, D-14059

Berlin, Germany. Web site http://www.rzpd.de

Seq primer: CGGTCGGGAATTCCTCCGGT.

FEATURES

source

1. .422

/organism="Mus musculus"

/strain="C57BL/6 X DBA"

/db_xref="taxon:10090"

/clone="528_13N20"

/clone_lib="Beddington mouse dissected endoderm"

/tissue_type="dissected endoderm"

/dev_stage="7.5 dpc"

/note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Cloned

unidirectionally. - High quality sequence only submitted. -

Average insert size: 1.2 kb (range: 0.2 - 2.kb)"

BASE COUNT 128 a 86 c 143 g 65 t

ORIGIN

alignment_scores:

Quality: 463.50 Length: 92
Ratio: 5.093 Gaps: 1
Percent Similarity: 98.913 Percent Identity: 98.913

alignment_block:

US-09-327-750D-35 x AL118340

Align seg 1/1 to: AL118340 from: 1 to: 422

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
146 ATGGCATCCAAATTTAAACAAGTCATCTGATCTCCTGCTGGAGAAAGA 195
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||
196 CAAAAAGACAAAAAGAGGTGGGAAGGCCCTCCAAACAAAGTGAAGAAGAC 245
|||||

6: 791-806, 1996.
TAG LIB-UI-R-BJO
TAG TISSUE=atrium at 16.5 dpc
TAG_SEQ=GATTC
BASE COUNT 146 a 120 c 94 g 180 t
ORIGIN
alignment_scores:
Quality: 459.00 Length: 97
Ratio: 5.044 Gaps: 0
Percent Similarity: 93.814 Percent Identity: 86.598
alignment_block:
US-09-327-750D-35 x AW251884/rev ..
Align seg 1/1 to reverse of: AW251884 from: 1 to: 540
21 ArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisHisLeuG1 37
:::|||||
538 AAAGGTGGGAGGCGCTCCAAACAAAGTGAAGAGGAATCCACCCTCGA 489
37 uGluValGluAsnLysLysProGlyGlyAsnValArgArgLysValArgA 54
|||||
488 AGAGGTTGAAACAAAGAGGCTGGGGCAATGTCAGGAGGAAAGTCAGGC 439
54 rgLeuValProAsnPheLeuTirPalaIleProAsnArgHisValAspArg 70
|||||
438 GACTGTGCTTAACCTTTCTATGGCCATACCTAATAGGCATGTTGATCAC 389
71 AsnGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluVa 87
:::|||||
388 AGTGAAGGGGAGAGGAGGTTGGAGATTCGTAGGCGAGGTGATGGAAGC 339
87 lLysArgLysThrThrGluGlnGlnValArgProTyrArgArgPheArgTf 104
|||||
338 CAAGAGAAAGTCTAAGGAGCAACAGATGAGGCGCTTACAGCGCTTCCGAA 289
104 hrProGluProAspAsnHisTyrAspPheCysLeuIlePro 117
|||||
288 CCCCTGAACCTGACAAATCATTACGACTTTTTCCTCATCACT 248

33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
246 CCCACCATCTGGAAGAGGTTGAAACAAAGAGGCTGGGGGAAATGTCGGA 295
50 ArgLysValArgArgLeuValProAsnPheLeuTirPalaIleProAsnAr 66
|||||
296 AGGAAAGTCAGGCGACTTGTGCTTAACCTTCTCTGGGCCATACCAATAG 345
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
346 GCATGTGTATCGCAATGAGGGGGAGAGGATGTTGGGAGATTTCGTGTC 395
83 lngLysThrGluValLysArgLysThr 91
|||||
396 AGGGAACAGAGTCAAGAGAAAGACT 421
seq_name: gb_est1.AW251884
seq_documentation_block:
LOCUS AW251884 540 bp mRNA EST 17-DEC-1999
DEFINITION UI-R-BJO-adn-b-06-0-UI.s1 UI-R-BJO Rattus norvegicus cDNA clone
UI-R-BJO-adn-b-06-0-UI 3', mRNA sequence.
ACCESSION AW251884
VERSION AW251884.1 GI:6595475
KEYWORDS EST.
SOURCE Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 540)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msaores@blue.weeg.uiowa.edu
The sequence contained an oligo-dt track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dt track served to identify it as a clone from the
normalized atrium at 16.5 dpc library cDNA Library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..540
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJO-adn-b-06-0-UI"
/clone_lib="UI-R-BJO"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified
polylinker; Site1: Not I; Site2: Eco RI; The UI-R-BJO
library is a subtracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dt track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research

OM of: US-09-327-750D-35 to: N_Geneseq_1101.* out_format : pfs
 Date: Mar 11, 2002 3:42 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODELframet_p2n.model -DEV-xlh
 -Q/c9n2_1/USPTO.spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
 -DB-N_Geneseq_1101 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blowsum62
 -TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09327750_@CGN1_1_330 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-35
 Query length: 117
 Database: N_Geneseq_1101.*
 Database sequences: 930621
 Database length: 428662619
 Search time (sec): 355.560000

score_list:

Sequence	Strd Orig	Zscore	Escore	Len	Documentation
/SID2/gcgdata/geneseq/NA2001.DAT:AAH03517	+	355.50	767.81	1.1e-34	865
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13750	+	355.50	764.27	1.7e-34	1229
/SID2/gcgdata/geneseq/NA2001.DAT:AAH85548	+	355.50	763.22	2.0e-34	1364
/SID2/gcgdata/geneseq/NA2001.DAT:AAH23529	+	175.00	367.98	2.0e-12	891
/SID2/gcgdata/geneseq/NA2000.DAT:AAH21748	+	174.00	365.47	2.8e-12	917
/SID2/gcgdata/geneseq/NA2001.DAT:AAH23528	+	153.00	324.71	7.0e-10	700
/SID2/gcgdata/geneseq/NA2000.DAT:AAH01005	+	147.00	312.12	2.6e-09	485
/SID2/gcgdata/geneseq/NA2000.DAT:AAH01004	+	141.00	297.91	1.6e-08	532
/SID2/gcgdata/geneseq/NA2001.DAT:AAH75810	+	99.50	202.37	0.0034	767
/SID2/gcgdata/geneseq/NA2001.DAT:AAH55881	+	99.50	201.19	0.0039	862
/SID2/gcgdata/geneseq/NA2001.DAT:AAH59611	+	99.50	200.78	0.0042	898
/SID2/gcgdata/geneseq/NA1999.DAT:AAH23519	+	98.50	159.28	0.8526	4453
/SID2/gcgdata/geneseq/NA2001.DAT:AAH45143	+	94.00	189.87	0.0169	792
/SID2/gcgdata/geneseq/NA2000.DAT:AAH01003	+	92.00	182.99	0.0124	401
/SID2/gcgdata/geneseq/NA2001.DAT:AAH20201	+	89.00	189.41	0.0179	276
/SID2/gcgdata/geneseq/NA2001.DAT:AAH45401	-	89.00	189.41	0.0179	276
/SID2/gcgdata/geneseq/NA2001.DAT:AAH05901	-	89.00	189.41	0.0179	276
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13308	+	87.50	181.99	0.0463	415
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13461	+	87.50	181.99	0.0463	415
/SID2/gcgdata/geneseq/NA2001.DAT:AAH03209	+	87.50	181.99	0.0463	415
/SID2/gcgdata/geneseq/NA2000.DAT:AAH03880	+	87.50	177.29	0.0847	662
/SID2/gcgdata/geneseq/NA2000.DAT:AAH03879	+	87.50	176.75	0.0907	698
/SID2/gcgdata/geneseq/NA2001.DAT:AAH160367	+	87.50	174.68	0.1184	858
/SID2/gcgdata/geneseq/NA2000.DAT:AAH08989	+	84.50	170.20	0.2101	692
/SID2/gcgdata/geneseq/NA2000.DAT:AAH70191	+	82.00	157.98	1.01	1344
/SID2/gcgdata/geneseq/NA1990.DAT:AAH06000	+	81.50	150.50	2.63	2533
/SID2/gcgdata/geneseq/NA2000.DAT:AAH64139	+	80.00	118.25	164.41	50000
/SID2/gcgdata/geneseq/NA1998.DAT:AAH21511	+	80.00	137.30	14.29	6755
/SID2/gcgdata/geneseq/NA2001.DAT:AAH71145	+	79.50	149.29	3.07	1839
/SID2/gcgdata/geneseq/NA2000.DAT:AAH3711	+	77.50	141.89	7.93	2471
/SID2/gcgdata/geneseq/NA2001.DAT:AAH13331	+	77.50	134.33	75.44	14141
/SID2/gcgdata/geneseq/NA2001.DAT:AAH29749	+	76.00	133.90	22.10	3929
/SID2/gcgdata/geneseq/NA1999.DAT:AAH13264	+	76.00	129.63	38.21	6005
/SID2/gcgdata/geneseq/NA1999.DAT:AAH231091	+	75.50	144.44	5.72	1236
/SID2/gcgdata/geneseq/NA1990.DAT:AAH06613	-	75.50	129.01	41.37	5721
/SID2/gcgdata/geneseq/NA1989.DAT:AAH91167	-	75.50	127.91	47.63	6381
/SID2/gcgdata/geneseq/NA2001.DAT:AAH22893	-	75.50	143.96	6.08	1161
/SID2/gcgdata/geneseq/NA2000.DAT:AAH38803	+	74.50	141.40	8.45	1342
/SID2/gcgdata/geneseq/NA2000.DAT:AAH49751	+	74.50	141.39	8.46	1343
/SID2/gcgdata/geneseq/NA1997.DAT:AAH12500	+	74.50	127.88	47.82	5137

/SID2/gcgdata/geneseq/NA1999.DAT:AAZ01022 - 74.50 103.74 1.1e+03 5652
 /SID2/gcgdata/geneseq/NA2001.DAT:AAI21493 + 74.00 152.63 2.00 394
 /SID2/gcgdata/geneseq/NA2001.DAT:AAI46784 + 74.00 152.63 2.00 394
 /SID2/gcgdata/geneseq/NA2001.DAT:AAI07190 + 74.00 152.63 2.00 394
 /SID2/gcgdata/geneseq/NA1995.DAT:AAQ86237 + 74.00 128.11 46.45 4500

seq_name: /SID2/gcgdata/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA: 865 BP.
 XX
 AC AAH03517;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:352.
 XX
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PR 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 XX
 PR 27-AUG-1999; 99JP-0300253.
 XX
 PR 11-JAN-2000; 2000JP-0118776.
 XX
 PR 02-MAY-2000; 2000JP-0183767.
 XX
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 (HELI-) HELIX RES INST.
 PA
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 WPI; 2001-318749/34.
 XX
 Claim 1: SEQ ID 352; 2537pp + CD ROM; English.
 XX
 The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesizing polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.

Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:
 Quality: 355.50 Length: 120
 Ratio: 3.665 Gaps: 3
 Percent Similarity: 80.833 Percent Identity: 59.167

alignment_block:
 US-09-327-750D-35 x AAH03517

Align seg 1/1 to: AAH03517 from: 1 to: 865

```

1 MetAlaSerLysLysGlnValIleLeu...AspLeuThrValGluLysAs 16
||| |||||.....: |||||: |||||:
191 ATGGATCCAAAGAGGAGTACTAGCGCAACATCTCAACGGGGAATGC 240

16 pLysLysAspLysArgGlyGly...LysAlaSerLysGlnSerGluGlu 32
: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
241 CCAACAAGAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 290

32 LuProHisHisLeuGluGluValGluAsnLysLysProGlyGlyVal 48
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 AATCCGCCCATTTGGGAGGGGTGAAGGCCAGAACGCTGGAGGAATATC 340

49 ArgArg...LysValArgArgLeuValProAsnPheLeuTirPAlaIlePr 64
||||| :|||: ||||| ||||| ||||| ||||| ||||| |||||
341 AGCGGGGGCGGAGTACTAGCGCAATCTCAACGGGGAATGC 240

64 OAsnArgHisValAspArgAsnGluGlyGluAspValGlyArgPheV 81
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 TANTAGGCATATTGAGCACAATGAGCGAGAGATGATGAGAAAGTTTG 440

81 alValGlnGlyThrGluValLysArgLysThrThrGluGlnValArg 97
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
441 TAGGCGAGATGATGGAATCAAGAGAAAGACTAGGGAACAGCAGATGAGG 490

98 ProTyrArgArgPheArgThrProGluProAspAsnHisTyrAspPheC 114
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 CACTATATGCGCTTCCAACTCTGAACTGACCTGACCACTATGACTTTTG 540

114 sLeuIlePro 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 CCTCATACCT 550

```

seq_name: /SIDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAH13750

seq_documentation_block:
 ID AAH13750 standard; cDNA: 1229 BP.

AC AAH13750;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:10656.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EPI074617-A2.

PD 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

PF 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs.

PS Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification, where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

CC AAH98993 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:
 Quality: 355.50 Length: 120
 Ratio: 3.665 Gaps: 3
 Percent Similarity: 80.833 Percent Identity: 59.167

alignment_block:
 US-09-327-750D-35 x AAH13750

Align seg 1/1 to: AAH13750 from: 1 to: 1229

```

1 MetAlaSerLysLysGlnValIleLeu...AspLeuThrValGluLysAs 16
||| |||||.....: |||||: |||||:
191 ATGGATCCAAAGAGGAGTACTAGCGCAACATCTCAACGGGGAATGC 240

16 pLysLysAspLysArgGlyGly...LysAlaSerLysGlnSerGluGlu 32
: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
241 CCAACAAGAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 290

32 LuProHisHisLeuGluGluValGluAsnLysLysProGlyGlyVal 48
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 AATCCGCCCATTTGGGAGGGGTGAAGGCCAGAACGCTGGAGGAATATC 340

49 ArgArg...LysValArgArgLeuValProAsnPheLeuTirPAlaIlePr 64
||||| :|||: ||||| ||||| ||||| ||||| ||||| |||||
341 AGCGGGGGCGGAGTACTAGCGCAATCTCAACGGGGAATGC 240

64 OAsnArgHisValAspArgAsnGluGlyGluAspValGlyArgPheV 81
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 TANTAGGCATATTGAGCACAATGAGCGAGAGATGATGAGAAAGTTTG 440

81 alValGlnGlyThrGluValLysArgLysThrThrGluGlnValArg 97
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
441 TAGGCGAGATGATGGAATCAAGAGAAAGACTAGGGAACAGCAGATGAGG 490

98 ProTyrArgArgPheArgThrProGluProAspAsnHisTyrAspPheC 114
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 CACTATATGCGCTTCCAACTCTGAACTGACCTGACCACTATGACTTTTG 540

```

```
114 sLeuIlePro 117
      |||||
541 CCTCATACCT 550
```

seq_name: /SIDS2/qcdata/qeneseq/qeneseq/NA2001.DAT:AAC85548

seq_documentation_block:
ID AAC85548 standard; cDNA: 1364 bp.

XX	AAC85548;	
AC		
XX		
DT	04-JUN-2001	(first entry)
XX		
DE	CDNA encoding CDIFF-4, Incvte ID No. 1990956CBL.	

XX	Human; cell differentiation; CDRF; agonist; antagonist; epilepsy;
XX	cell proliferation; Alzheimer's disease; schizophrenic disorder;
KW	arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.
XX	Human sapiens.
OS	Homo sapiens.

Key	Location/Qualifiers
CDS	250..612
	/*tag= a
	/product= "CDIFF-4"

PN WO200119860-A2.

22-MAR-2001

14-SEP-2000: 2000WO-US25435.

15-SEP-1999: 99US-0154140.

06-DEC-1999; 99US-0169155; PR

PA (INCY-) INCYTE GENOMICS INC.

PPI Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
 PI Baughn MR, Lu DAN, Bandman O, Shih LL, Patterson C;
 XX WPI; 2001-211447/21.
 DR P-PSDB: AAB47126.
 DR

Isolated polypeptides and polynucleotides involved in cell differentiation are used for treatment, prevention and diagnosis of cell proliferative, developmental and neurological disorders e.g. cancer and Alzheimer's disease.

PS
xx
Claim 5: page 121: 137pp: English.

The sequences given in AAC85545-72 encode human polypeptides involved in cell differentiation (CDIFF). CDIFF polypeptides and agonists of these are used to treat a disease or condition associated with decreased expression of functional CDIFF. An antagonist of CDIFF is used to treat a disease or condition associated with over expression of functional CDIFF. CDIFF polypeptides may be used for the treatment, prevention and diagnosis of cell proliferative, developmental and neurological disorders, such as Alzheimer's disease, schizophrenic disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus RX3-2. This sequence maps to chromosome 1 within the interval from 152.2 to 157.4 centiMorgans, to chromosome 3 within the interval from 157.4 to 158.0 centiMorgans, and to the x chromosome within the interval from 104.9 to 150.3 centiMorgans.

SQ Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:	
Quality:	355.50
Ratio:	3.665
Percent Similarity:	80.833
Percent Identity:	59.167
Gaps:	3
Length:	120

alignment_block:

US-09-327-750D-35 x AAC85548

Align seq 1/1 to: AAC85548 from: 1 to: 1364

seq_name: /SIDS2/qcadata/qeneseq/qeneseq/NA2001.DAT:AAF23529

seq_documentation_block:

ID: AAF23529 standard: DNA: 891 BP.

AC AAF23529:

DT 22-MAR-2001 (first entry)

Human NADE DNA-

KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.

AA Homo sapiens

PN WO2000075273-A2

14-DEC-2000

07 - JUN - 2000 - 2000W0 - UST 5621

XX
PP
07-JUN-1999
99JIS-0327750

XX
PA
UNYCO \ UNIV COLUMBYA NEW YORK

EX DT
E C
+ S
C

XX
WDT. 2001-051707/07

XX New p75-neurotrophin receptor-associated cell death executor (NADE), and
PT the gene encoding NADE, useful for modulating the activity of p75NTR
PT and for detecting neurodegenerative diseases -
XX
XX
PS Disclosure: Fig 1: 134pp; English.


```
alignment_scores:
  Quality: 99.50      Length: 132
  Ratio: 1.363       Gaps: 6
  Percent Similarity: 55.303   Percent Identity: 29.545

alignment_block:
US-09-327-750D-35 x AAF59611

Align seg 1/1 to: AAF59611 from: 1 to: 862
1 MetAlaSerLysLysGln.....ValIleLeuAspLe 11
226 ATGGAGTCCAAAGAGGAAGCAGCGTTAAACAATCTCATCGTGGAATAATGT 275
11 uThrValGluLysAspLysLysAspLysArgGlyGlyLysAlaSerLysG 28
276 CAACAGGAAATGATGATAAAGATGAAGAGGACCAAGTGTCTAATAAA. 324
28 InSerGluGluProHisHisLeuGluValGluAsnLysLysPro 44
325 ..GGGAGCCCTTGGCCCTTGAATGTTAGTGAATACTGTGTGCCT 372
45 GlyGlyAsnValArgArg...LysValArgArgLeuValProAsnPheLe 60
373 AGAGGAAACCGTAGCGGTCGCGTTAGCAGCCCACTCTGCAGTATAG 422
60 uTrpAlaIleProAsnArgHisValAsp.....ArgAsnG 72
423 ATGGGACATAATGATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 472
72 LuGlyGlyGluAspValGlyArgPheValGlnGlyThrGluValLys 88
473 AGAATATGAAAGGATTCGGAGAGGTGACACAG.....CTGATG 513
89 ArgLysThrThrGluGlnValArgProTyrArgArgPheArgThrPr 105
514 GAAAGCTGAGGAAAGCAGTTGATGATAGTTTGGCGGCAGTCAGCAC 563
105 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 117
564 TGATCCCTCCATCCATGACCATCAGGATGTTTGCCTTATGCC 609
seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF59611
seq_documentation_block:
ID AAF59611 standard; cDNA; 898 BP.
XX AAF59611;
AC AAF59611;
DT 24-APR-2001 (first entry)
DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.
KW Cell cycle and proliferation protein; CCYPR; human; agonist;
KW antagonist; gene therapy; detection; gene therapy;
KW transgenic animal disease model; immune disorder;
KW developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX Homo sapiens.
XX WO200107471-A2.
XX 01-FEB-2001.
XX 21-JUL-2000; 2000WO-US19948.
XX 21-JUL-1999; 99US-0145075.
XX 08-SEP-1999; 99US-0153129.
PR
```

```
PR 10-NOV-1999; 99US-0164647.
XX (INCY-) INCYTE GENOMICS INC.
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzal Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX WPI; 2001-112727/12.
DR P-PSDB; AAB60474.
XX Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signaling disorders and cell proliferative disorders including cancer -
XX Claim 5; Page 181-182; 205pp; English.
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;
SQ
```

```
alignment_scores:
  Quality: 99.50      Length: 132
  Ratio: 1.363       Gaps: 6
  Percent Similarity: 55.303   Percent Identity: 29.545

alignment_block:
US-09-327-750D-35 x AAF59611

Align seg 1/1 to: AAF59611 from: 1 to: 898
1 MetAlaSerLysLysGln.....ValIleLeuAspLe 11
254 ATGGAGTCCAAAGAGGAAGCAGCGTTAAACAATCTCATCGTGGAATAATGT 303
11 uThrValGluLysAspLysLysAspLysArgGlyGlyLysAlaSerLysG 28
304 CAACAGGAAATGATGATAAAGATGAAGAGGACCAAGTGTCTAATAAA. 352
28 InSerGluGluProHisHisLeuGluValGluAsnLysLysPro 44
353 ..GGGAGCCCTTGGCCCTTGAATGTTAGTGAATACTGTGTGCCT 400
45 GlyGlyAsnValArgArg...LysValArgArgLeuValProAsnPheLe 60
401 AGAGGAAACCGTAGCGGTCGCGTTAGCAGCCCACTCTGCAGTATAG 450
60 uTrpAlaIleProAsnArgHisValAsp.....ArgAsnG 72
451 ATGGGACATAATGATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGG 500
72 LuGlyGlyGluAspValGlyArgPheValGlnGlyThrGluValLys 88
```

```

||||| ||| ..... ||| ||| .....
501 AGAATATGGAAGATTGGGAGGAGTGACAG.....CTGATG 541
89 ArglyeThrThrGluGlnValArgProTyrArgPheArgThrPr 105
   ||||| ..... ||||| .....
542 GAAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 591
105 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 117
   ||||| ..... ||||| ..... ||||| ..... ||||| .....
592 TGATCCCTCCATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 637

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAx23519

seq_documentation_block:
ID AAX23519 standard; DNA; 44453 BP.

AC AAX23519;
DT 23-JUN-1999 (first entry)
DE Human kidney aminopeptidase P genomic DNA fragment 3.

XX Aminopeptidase: human; Amp; gene therapy; treatment; Amp-deficiency;
KW prenatal diagnosis; angiodema; antihypertensive agent; atherosclerosis;
KW arterial stenosis; industrial protein feed; malabsorption syndrome;
KW proteinaceous waste degradation; additive; immunohistochemistry; ss.

XX Homo sapiens.
OS Wo9911799-A2.

XX 11-MAR-1999.

XX 02-SEP-1998; 98WO-US18426.

XX 02-SEP-1997; 97US-0057854.

XX (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.

XX Ryan JW, Sprinkle TJC, Venema RC;

XX WPI; 1999-205193/17.

XX Nucleic acid encoding human aminopeptidase P

XX Claim 13; Page 139-165; 201pp; English.

XX This invention describes the isolation of a novel human aminopeptidase P
CC (Amp). This protein is used to produce recombinant Amp and can be used
CC for gene therapy for treating Amp-deficiency conditions. Its fragments
CC are used as primers and probes to identify patients with homozygous and
CC heterozygous Amp deficiency, including prenatal diagnosis (patients
CC defective in Amp are at risk of developing angiodema if treated with
CC angiotensin-converting enzyme inhibitors), also as antisense inhibitors
CC in cases of excessive Amp expression. The product of the invention is
CC also used to identify Amp-expressing sequences in other animals and to
CC generate transgenic animals, and comparisons of genomic sequences are
CC used to detect mutations. Amp inhibitors are potentially useful as
CC antihypertensive agents and to prevent or treat arterial (re)stenosis
CC or atherosclerosis. The structure of Amp is used to design synthetic
CC substrates, e.g. for use in Amp assays. Amp, which hydrolyzes N-terminal
CC imido bonds, can be used to degrade industrial protein feeds to free
CC amino acids, to degrade proteinaceous wastes, as additives in enzyme
CC formulations used to treat malabsorption syndrome and for studying its
CC biological role. Antibodies against Amp are used in immunohistochemical
CC methods to study Amp distribution.

XX Sequence 44453 BP; 13034 A; 10110 C; 10292 G; 11017 T; 0 other;

alignment_scores:
Quality: 98.50 Length: 102
Ratio: 1.539 Gaps: 2

Percent Similarity: 62.745 Percent Identity: 26.471

alignment_block:

US-09-327-750D-35 x AAX23519

Align seg 1/1 to: AAX23519 from: 1 to: 44453

```

1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAspLys 17
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
43713 CTACAAGAAGAGGAGGACACTGGCTTTTCAGTCAGTCAGCTTGGAGGAGAGAG 43762
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
43763 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 43812
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
34 1sHisLeuGluGluValGluAsnLysLysProGly..GlyAsnValArgA 50
   ||||| ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
43813 TGGACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 43862
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
50 rGlyValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArg 66
   ||||| ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
43863 GGAAGAAGAAGAGG.....AAGTGGAGGAGGCCAGAGAAAG 43897
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
67 HisValAspArgAsnGluGlyGluAspValGlyArgPheValValGlu 83
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
43898 GAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 43947
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
83 nGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrA 100
   ||||| ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
43948 AGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 43997
   ..... ||||| ..... ||||| ..... ||||| ..... ||||| .....
100 rGArg 101
   ..... ||||| .....
43998 GCAAA 44002

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard; cDNA; 792 BP.

XX AAH45143;

XX 07-SEP-2001 (first entry)

XX Human brain expressed X-linked protein, hBex, coding sequence.

XX Human: brain expressed X-linked protein; cytostatic; auditory; nootropic;
KW hBex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;
KW X-chromosome-binding mental retardation; lissencephalous disease; ss.

XX Homo sapiens.

XX WO200140286-A1.

XX 07-JUN-2001.

XX 27-NOV-2000; 2000WO-CN00502.

XX 30-NOV-1999; 99CN-0124179.

XX (BIOR-) BIORAD GENE DEV LTD SHANGHAI.

XX Mao Y, Xie Y;

XX WPI; 2001-397944/42.

XX P-PSDB; AAB99224.

XX Isolated human brain-expressed X-linked polypeptide used to diagnose
PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
PT deafness and X-chromosome-binding mental retardation

XX Claim 5; Page 22; 30pp; Chinese.


```
KW Probe: human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX OS Homo sapiens.
XX PN W0200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234697.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000US-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX PS Claim 25; SEQ ID NO 10134; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX (SENP). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX CC Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 276 BP; 4 A; 131 C; 5 G; 136 T; 0 other;

alignment_scores:
  Quality: 89.00 Length: 94
  Ratio: 1.712 Gaps: 3
  Percent Similarity: 55.319 Percent Identity: 31.915

alignment_block:
US-09-327-750D-35 x AA120201/rev ..

Align seg 1/1 to reverse of: AA120201 from: 1 to: 276

14 GluLysAspLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerG1 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 GAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 223

30 uGluGluProHisHLeuGluGluValGluAsnLysLysProGlyGlyA 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 173

47 snValArgArgLysValArgLeuValProAsnPheLeuTrpAlaIle 63
:::|||||:|||||:|||||:|||||:|||||:|||||:
172 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150

64 ProAsnArgHisValAspArgAsnGluGly.Gly.....G 75
:::|||||:|||||:|||||:|||||:|||||:
149 GAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 100

75 LuAspValGlyArgPheValValGlnGlyThrGluValLysArgLysThr 91
|||||:|||||:|||||:|||||:|||||:|||||:
99 AGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 50

92 ThrGluGlnGlnValArgProTyrArgArg 101
|||||:|||||:|||||:|||||:
49 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 20
```



```
LOCATION: 13617..13636
FEATURE:
OTHER INFORMATION: downstream amplification primer 4-72, complement
NAME/KEY: primer_bind
LOCATION: 13547..13564
OTHER INFORMATION: upstream amplification primer 4-73
FEATURE:
NAME/KEY: primer_bind
LOCATION: 13962..13981
OTHER INFORMATION: downstream amplification primer 4-73, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 15994..16011
OTHER INFORMATION: downstream amplification primer 99-610
FEATURE:
NAME/KEY: primer_bind
LOCATION: 16463..16480
OTHER INFORMATION: upstream amplification primer 99-610, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 17304..17324
OTHER INFORMATION: downstream amplification primer 99-609
FEATURE:
NAME/KEY: primer_bind
LOCATION: 17814..17832
OTHER INFORMATION: upstream amplification primer 99-609, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 18008..18027
OTHER INFORMATION: upstream amplification primer 4-90
FEATURE:
NAME/KEY: primer_bind
LOCATION: 18423..18442
OTHER INFORMATION: downstream amplification primer 4-90, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 18699..18716
OTHER INFORMATION: downstream amplification primer 99-607
FEATURE:
NAME/KEY: primer_bind
LOCATION: 19164..19182
OTHER INFORMATION: upstream amplification primer 99-607, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 22589..22609
OTHER INFORMATION: downstream amplification primer 99-602
FEATURE:
NAME/KEY: primer_bind
LOCATION: 23111..23129
OTHER INFORMATION: upstream amplification primer 99-602, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 25098..25118
OTHER INFORMATION: downstream amplification primer 99-600
FEATURE:
NAME/KEY: primer_bind
LOCATION: 25657..25674
OTHER INFORMATION: upstream amplification primer 99-600, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 26537..26557
OTHER INFORMATION: downstream amplification primer 99-598
FEATURE:
NAME/KEY: primer_bind
LOCATION: 27022..27040
OTHER INFORMATION: upstream amplification primer 99-598, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 32262..32281
OTHER INFORMATION: downstream amplification primer 99-592
FEATURE:
NAME/KEY: primer_bind
LOCATION: 32823..32841
```

```
OTHER INFORMATION: upstream amplification primer 99-592, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 34215..34233
OTHER INFORMATION: upstream amplification primer 99-217
FEATURE:
alignment_scores:
Quality: 74.50 Length: 82
Ratio: 1.656 Gaps: 2
Percent Similarity: 54.878 Percent Identity: 30.488
alignment_block:
US-09-327-750D-35 x US-09-338-907-179/rev ..
Align seg 1/1 to reverse of: US-09-338-907-179 from: 1 to: 56520
14 GlutysapLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGl 30
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39085 GAAAGCGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 39036
30 uGluCluProHisHisLeuGluGluValGluAsnLysLysProGlyGlyA 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39035 AACGAGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 38986
47 snValArgArgLysValArgLysValProAsnPheLeuTrpAlaIle 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
38985 GAGAGAGAGCGAAGAAAGAGAGG..... 38963
64 ProAsnArgHisValAspArgAsnGluGlyGlyGlu.....As 76
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
38962 ..... 38931
76 pValcArgPheValValGlnGlyThrGluValLysArgLysThr 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
38930 AATTGGGCATCTAGGCTGAAGGACAAACATCTAGAGTAGT 38885
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-996-306-1
seq_documentation_block:
Sequence 1, Application US/08996306
Patent No. 5945522
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Chumakov, Ilya
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: prostate cancer gene
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSER: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,306
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: GENSET.018A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1:
```

```

; NAME/KEY: SEQ ID65
; LOCATION: 51149...51168
; FEATURE:
; NAME/KEY: SEQ ID59
; LOCATION: 51448...51494
; FEATURE:
; NAME/KEY: SEQ ID68
; LOCATION: compl(51482..51499)
; FEATURE:
; NAME/KEY: SEQ ID44
; LOCATION: 51596...51613
; FEATURE:
; NAME/KEY: SEQ ID26
; LOCATION: 51612...51658
; FEATURE:
; NAME/KEY: SEQ ID53
; LOCATION: compl(51996..52015)
; FEATURE:
; NAME/KEY: polyAD signal
; LOCATION: 54445..54450
;
US-08-996-306-1

alignment_scores:
Quality: 71.50 Length: 79
Ratio: 1.663 Gaps: 2
Percent Similarity: 54.430 Percent Identity: 30.380

alignment_block:
US-09-327-750B-35 x US-08-996-306-1/rev ..
Align seg 1/1 to reverse of: US-08-996-306-1 from: 1 to: 56516

17 LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluPr 33
||||| :||||| :||||| :||||| :||||| :||||| :||||| :
39077 AACAAGAGAGAGAGAGAGAGAGAGCAACGAAGCAAGCAAGCAAGTGA 39028
33 oHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArgA 50
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
39027 GCGAAGCAAGAAAGGAGGAGGGAAGAAAGAAAGAAAGAGGAGAGAGAG 38978
50 rGlyValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArg 66
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
38977 CGAAGCAAGAAAGG..... 38964
67 HisValAspArgAsnGluGlyGlyGlu.....AspValGlyArg 79
: : : : : : : : : : : : : : : : : : : : : : : : : :
38963 .....AAGAAAGAGGGGAGAGAGCAATGAGTTCAAATTGGGCA 38923
79 gPheValValGlnGlyThrGluValLysArgLysThr 91
: : : : : : : : : : : : : : : : : : : : : : : : : :
38922 TACTGAGGCTGAAGGAACAAACATTTCTAGAGTAGT 38886

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-338-907-1

seq_documentation_block:
; Sequence 1, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338-907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996.306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099.658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218.207
; EARLIER FILING DATE: 1998-12-22

```



```

; SEQ ID NO 5
; LENGTH: 4094
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (126)..(3479)
US-08-841-483-5

```

```

alignment_scores:
  Quality: 71.00      Length: 82
  Ratio: 1.690        Gaps: 5
  Percent Similarity: 51.220  Percent Identity: 31.707

```

```
alignment_block:
```

```
US-09-327-750D-35 x US-08-841-483-5 ..
```

```
Align seg 1/1 to: US-08-841-483-5 from: 1 to: 4094
```

```

12 ThrValGluLysAspLysLysAspLysArgGlyGly.....LysAL 25
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1452 ACTCTGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGTC 1501
25 aserLysGlnSerGluGluGlu..... 32
1502 CAGCAAGAAAGGCGCTGAGGAGGCGCTGGAGACCCCTTCATCATCAGGC 1551
33 .....ProHisHisLeuGluGlu.....ValGluAsnLysLys 43
1552 CCACCCCGCTCCCGCTCATGAAGCCCTGCTGGTGTGTTGTGAACCCCAAG 1601
44 ProGlyGlyAsnValArgArgLysValArgLeuValProAsnPhLe 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1602 AGTGGGGCAACCAAGGTCGAAGATC.....ATCCAGTCTTCT 1642
60 utrPalalleProAsnArgHisVal....AspArgAsnGluGlyGly 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1643 CTGGTATCTCAATCCCGCAAGTCTTCGACCTGAGCCAGGAGGG 1688
seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-382-911-5

```

```

seq_documentation_block:
; Sequence 5, Application US/09382911
; Patent No. 6221658
; GENERAL INFORMATION:
; APPLICANT: Prescott, Steven M.
; APPLICANT: Bunting, Michaeline
; APPLICANT: Tang, Wen
; APPLICANT: Topham, Matthew
; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
; FILE REFERENCE: 2037.2.1s
; CURRENT APPLICATION NUMBER: US/09/382,911
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 08/841,483
; PRIOR FILING DATE: 1997-04-22
; PRIOR APPLICATION NUMBER: 60/016,210
; PRIOR FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4094
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (126)..(3479)
US-09-382-911-5

```

```

alignment_scores:
  Quality: 71.00      Length: 82

```

```

Ratio: 1.690        Gaps: 5
Percent Similarity: 51.220  Percent Identity: 31.707
alignment_block:
US-09-327-750D-35 x US-09-382-911-5 ..
Align seg 1/1 to: US-09-382-911-5 from: 1 to: 4094
12 ThrValGluLysAspLysLysAspLysArgGlyGly.....LysAL 25
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1452 ACTCTGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGTC 1501
25 aserLysGlnSerGluGluGlu..... 32
1502 CAGCAAGAAAGGCGCTGAGGAGGCGCTGGAGACCCCTTCATCATCAGGC 1551
33 .....ProHisHisLeuGluGlu.....ValGluAsnLysLys 43
1552 CCACCCCGCTCCCGCTCATGAAGCCCTGCTGGTGTGTTGTGAACCCCAAG 1601
44 ProGlyGlyAsnValArgArgLysValArgLeuValProAsnPhLe 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1602 AGTGGGGCAACCAAGGTCGAAGATC.....ATCCAGTCTTCT 1642
60 utrPalalleProAsnArgHisVal....AspArgAsnGluGlyGly 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1643 CTGGTATCTCAATCCCGCAAGTCTTCGACCTGAGCCAGGAGGG 1688
seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-402-282-3

```

```

seq_documentation_block:
; Sequence 3, Application US/08402282
; Patent No. 5476768
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Beyer Jr., Wayne E.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,282
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 222..425

```

```

; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function- "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product- "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
; OTHER INFORMATION: /function- "potential open reading
; OTHER INFORMATION: /frame"
; US-08-402-282-3
;
alignment_scores:
; Quality: 71.00 Length: 71
; Ratio: 1.775 Gaps: 4
; Percent Similarity: 56.338 Percent Identity: 33.803
;
alignment_block:
; US-09-327-750D-35 x US-08-402-282-3
;
Align seg 1/1 to: US-08-402-282-3 from: 1 to: 15664
;
49 ArgArgLysValArgLeuValProAsnPheLeuTrpAlaIleProAs 65
|||||
13436 CGAAATCGGTGGTGGCAAGGCGCGCGATGCGCGTGG.....CCGAC 13479
;
65 narg.His.....ValAspArgAsnGluGly 73
|||||
13480 GCGCTCACCCCAACCGCGACTCGAGTCCGGTCGATCGGTCCGGGA 13529
;
74 GlyGluAspValGlyArgPheValGlnGlyThrGluValLysArgLy 90
|||||
13530 GGGATGAT.....CGAAGCTTACCGCGCGGACACGACGAAACGGCG 13573
;
90 sThrThrGluGlnValArgProTyrArgPheArgThrProGluP 107
|||||
13574 GCGACGACGACGGAATCGACGCGCCACGACACGATCGTGACGTGATC 13623
;

```

```
107 roaspsnHls 110
|| :|||||
13624 CTTTCAGCAT 13634

seq_name: /cgn2_6/ptodata/2/ins/5A_COMB.seq:US-08-508-004-3

seq_documentation_block:
: Sequence 3, Application US/08508004
: Patent No. 5582969
: GENERAL INFORMATION:
: APPLICANT: Pearson, Robert E.
: APPLICANT: Dickson, Julie A.
: APPLICANT: Hamilton, Paul T.
: APPLICANT: Little, Michael C.
: APPLICANT: Beyer Jr., Wayne F.
: TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
: TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
: ADDRESSEE: Company
: STREET: 1 Becton Drive
: CITY: Franklin Lakes
: STATE: NJ
: COUNTRY: US
: ZIP: 07417
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/508,004
: FILING DATE: 27-JUL-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/402,282
: FILING DATE: 10-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Fugit, Donna R.
: REGISTRATION NUMBER: 32,135
: REFERENCE/DOCKET NUMBER: P-3283
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 15864 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 222..425
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 451..747
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 747..1109
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1109..2014
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2034..2747
```

```
:
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2747..3109
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3109..3444
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3444..3728
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3731..4855
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 4855..5376
: OTHER INFORMATION: /function= "potential coding
: OTHER INFORMATION: /product= "L5 gp37 homolog"
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 5382..5747
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 5837..6307
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 6403..7770
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 7770..8006
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 8033..8236
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 8244..9443
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 9450..10244
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 10371..10586
: OTHER INFORMATION: /function= "potential open reading
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 11115..11786
: OTHER INFORMATION: /function= "potential open reading
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14771..15154
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15154..15426
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15429..15664
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
US-08-508-004-3
alignment_scores:
Quality: 71.00 Length: 71
Ratio: 1.775 Gaps: 4
Percent Similarity: 56.338 Percent Identity: 33.803
alignment_block:
US-09-327-750D-35 x US-08-508-004-3 ..
Align seg 1/1 to: US-08-508-004-3 from: 1 to: 15664
49 ArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAs 65
|||||
13436 CGAATCGGTGGCGGCAAGGGCGCGGATGCCGTGG.....CCGAC 13479
|||||
65 Arg.His.....ValAspArgAsnGluGly 73
|||||
13480 CGGCTCACCCGCAACCCGACCTCGAGTCCGTCGATCGTCCGGGA 13529
|||||
74 GlyGluAspValGlyArgPheValValGlnGlyThrGluValLysArgly 90
|||||
13530 GGGATGAT.....CGAGCCTACCGCGGACACGAGAGAAACGGCG 13573
|||||
90 sThrThrGluGlnValArgProTyrArgPheArgPheArgThrProGluP 107
|||||
13574 GCGAGCGAGCGGAATCGACGGCCGACAGACAGCATCGTGCATCGATC 13623
|||||
107 roAspAsnHis 110
|||
13624 CTGTTCCAGCAT 13634
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-066-3
seq_documentation_block:
Sequence 3, Application US/08402066
Patent No. 5612182
GENERAL INFORMATION:
APPLICANT: Pearson, Robert E.
APPLICANT: Dickson, Julie A.
APPLICANT: Hamilton, Paul T.
APPLICANT: Little, Michael C.
APPLICANT: Beyer Jr., Wayne F.
TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
ADDRESSEE: Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/402,066
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3283
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 15664 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 222..425
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 451..747
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 747..1109
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1109..2014
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2034..2747
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2747..3109
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3109..3444
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3444..3728
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3731..4855
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
```

```

NAME/KEY: misc_feature
LOCATION: 4855..5376
OTHER INFORMATION: /function= "potential coding"
OTHER INFORMATION: sequence
OTHER INFORMATION: /product= "L5 gp37 homolog"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5382..5747
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5837..6307
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6403..7770
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7770..8006
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8033..8236
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8244..9443
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9450..10244
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10371..10586
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11115..11786
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14771..15154
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15154..15426
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15429..15664

```

```

; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
US-08-402-066-3

alignment_scores:
Quality: 71.00 Length: 71
Ratio: 1.775 Gaps: 4
Percent Similarity: 56.338 Percent Identity: 33.803

alignment_block:
US-09-327-750D-35 x US-08-402-066-3
Align seg 1/1 to, US-08-402-066-3 from: 1 to: 15664
49 ArgArgLysValArgLeuValProAsnPheLeuTrpAlaIleProAs 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CGAATCGGTGGCGGCAAGGCGCGCGATGGCGGTGG.....CCGAC 13479
65 nArg.His'.....ValasPArgAsnGluGly 73
|||||:|||||:|||||:|||||:|||||:|||||:
13480 CGCCTCACCCCGAACCGGACCTCGAGGTCCGGTCGATCGTCCGGGA 13529
74 GlyGluAspValGlyArgPheValValGlnGlyThrGluValLysArgLy 90
|||||:|||||:|||||:|||||:|||||:|||||:
13530 GGGATGAT.....CGAAGCCTACCGGGGCGGACACGAAGAAACGGCG 13573
90 sThrThrGluGlnValArgProTyrArgArgPheArgThrProGluP 107
|||||:|||||:|||||:|||||:|||||:|||||:
13574 GCGACGACGACGGAATCGACGGCCGACGACGATCGTCGATCGTCGATC 13623
107 roAspAsnHis 110
||:|||||
13624 CTGTTCCAGCAT 13634

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-068-3

seq_documentation_block:
; Sequence 3, Application US/08402068
; Patent No. 5633159
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Beyer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,068
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs

```



```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 222..425
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 451..747
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 747..1109
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1109..2014
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2034..2747
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2747..3109
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3109..3444
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3444..3728
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3731..4855
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4855..5376
OTHER INFORMATION: /function= "potential coding
OTHER INFORMATION: sequence"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5382..5747
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5837..6307
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6403..7770
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7770..8006

```

```

OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8033..8236
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8244..9443
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9450..10234
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10371..10586
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11115..11786
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14771..15154
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15154..15426
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15429..15664
OTHER INFORMATION: /function= "potential open reading
OTHER INFORMATION: frame"
US-08-402-068-3
alignment_scores:
Quality: 71.00 Length: 71
Ratio: 1.775 Gaps: 4
Percent Similarity: 56.338 Percent Identity: 33.803
alignment_block:
US-09-327-750D-35 x US-08-402-068-3
Align seg 1/1 to: US-08-402-068-3 from: 1 to: 15664
49 ArgArgLysValArgLeuValProAsnPheLeuTtpAlaIleProAs 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CGAATCGGTGGCGGCAAGGGCGCCCGATGGCGCTGG.....CCGAC 13479
65 nArg.His.....ValaspArgAsnGluGly 73
|||||:|||||:|||||:|||||:|||||:|||||:
13480 GCGCTCACCCCGAACCGCACCTCGAGTCCGATCGATCGGTTCGGGGA 13529

```


OM of: US-09-327-750D-35 to: GenEmbl.* out_format : pfs

Date: Mar 11, 2002 3:35 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n_model -DEV=xlh
-O=/cgn2_1/USFTO_spool/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
-DB=GenEmbl -qfmt=fastap -suffix=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELAEXT=7.000 -XGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTPMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750.ecgn1_1.8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-35

Query length: 117

Database: GenEmbl.*

Database sequences: 1472140

Database length: -341344837

Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_pr:AK000959	+	355.50	626.28	1.3e-26	1229	AK000959 Homo sapiens CDNA FLJ
gb_pa:AX100231	+	355.50	625.44	1.5e-26	1364	AX100231 Sequence 32 from Pat
gb_pr:HS635G19	+	355.50	593.78	8.5e-25	69648	AL0035494 Human DNA sequence f
gb_pr:BC003190	+	175.00	306.91	8.1e-09	793	BC003190 Homo sapiens, p75NTR-a
gb_pr:HMUGC	+	175.00	305.97	9.1e-09	891	M38188 Human unknown protein fr
gb_pr:AF187064	+	175.00	305.97	9.1e-09	891	AF187064 Homo sapiens p75NTR-as
gb_ro:AF187066	+	154.00	270.34	8.8e-07	700	AF187066 Mus musculus p75NTR-as
gb_ro:AF097440	+	150.00	268.74	1.1e-06	854	AF097440 Mus musculus brain exp
gb_sts:G72708	+	150.00	267.28	1.3e-06	421	G72708 MARC 4953-4954:99199031
gb_ro:AF187065	+	149.50	264.70	1.8e-06	519	AF187065 Rattus norvegicus p75N
gb_pr:HSV351F8	+	138.00	208.08	0.0026	45678	G720719 Human DNA sequence fr
gb_sts:G35294	+	120.50	213.50	0.0013	477	G35294 Human STS SHGC-37409, se
gb_pr:HS71487	+	114.50	159.87	1.25	98274	AF097438 Mus musculus brain exp
gb_ro:AF097438	+	107.50	185.74	0.0454	835	AF097438 Mus musculus brain exp
gb_ro:AF051347	+	107.50	185.71	0.0456	838	AF051347 Mus musculus REX-3 mR
gb_ro:AF097437	+	107.50	177.69	0.1275	2269	AF097437 Mus musculus Bex1 pr
gb_ro:AF097439	+	102.00	176.39	0.1506	785	AF097439 Mus musculus brain exp
gb_pr:AF251053	+	99.50	172.35	0.2528	744	AF251053 Homo sapiens X-linked
gb_pa:AC008272	+	99.50	170.84	0.3070	898	AC008272 Sequence 76 from Paten
gb_pr:HSV870H8	+	99.50	142.25	12.02	31321	AF20333 Human DNA sequence fr
gb_pr:AL133348	+	99.50	140.16	15.70	40584	AL133348 Human DNA sequence f
gb_htg:HS75399	+	99.50	140.07	15.88	41029	AL022169 Homo sapiens Chromos
gb_pr:HS75399	+	98.50	129.88	58.71	116561	AL023653 Human DNA sequence
gb_pr:AF195953	+	98.50	125.19	107.09	208618	AF195953 Homo sapiens membra
gb_htg:AC012947	+	95.00	121.94	162.57	143599	AC012947 Drosophila melanoga
gb_in:AC009183	+	95.00	119.60	219.49	192055	AC009183 Drosophila melanoga
gb_in:AE003685	+	95.00	118.32	221.08	193398	AE003685 Drosophila melanoga
gb_pr:AF183416	+	94.00	163.70	0.7668	692	AF183416 Homo sapiens ovarian g
gb_pr:AF237783	+	94.00	162.02	0.9513	791	AF237783 Homo sapiens brain-exp
gb_pr:AF220189	+	94.00	161.65	0.9972	828	AF220189 Homo sapiens uncharact
gb_pr:HS198P4	+	94.00	131.35	48.61	35714	AL008708 Human DNA sequence f
gb_pr:AC002478	+	91.00	120.55	194.16	70113	AC002478 Human BAC clone GSI-
gb_htg:AC012232	+	91.00	112.22	497.28	174337	AC012232 Homo sapiens clone
gb_da:WAFR2	+	90.50	146.47	6.99	2507	J04157 M.xanthus fr2A, fr2B an
gb_pr:AP0000265	+	90.50	124.52	116.65	38304	AP0000265 Homo sapiens genomic
gb_pr:AP000034	+	90.50	116.80	314.19	100000	AP000034 Homo sapiens genomic
gb_pr:AP000101	+	90.50	116.80	314.19	100000	AP000101 Homo sapiens genomic
gb_pr:AP000177	+	90.50	116.80	314.19	100000	AP000177 Homo sapiens genomic
gb_htg:AC026930	+	90.50	114.66	413.39	130443	AC026930 Homo sapiens chromo

gb_pr:AC005295 + 90.50 114.62 415.44 131070 ! AC005295 Homo sapiens xp2
gb_pr:AP001713 + 90.50 106.95 1.1e+03 340000 ! AP001713 Homo sapiens gen
gb_da:AE008371 - 90.00 133.00 39.36 11967 ! AE008371 Agrobacterium tum
gb_pr:U72787 + 90.00 129.48 61.75 18510 ! U72787 Homo sapiens cosmid
gb_pr:AL353806 + 90.00 115.06 392.66 111051 ! AL353806 Human DNA sequen

seq_name: gb_pr:AK000959

seq_documentation_block:
LOCUS AK000959 1229 bp mRNA PRI 22-FEB-2000
DEFINITION Homo sapiens CDNA FLJ10097 fls, clone HEMBA1002458, weakly similar
to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.
ACCESSION AK000959
VERSION AK000959.1 GI:7021945
KEYWORDS oligo capping; fls (full insert sequence).
Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to
SOURCE MRNA, clone lib:HEMBA1 clone:HEMBA1002458.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Negahari,K., Masuno,Y. and Oshima,A.
NEDO human cDNA sequencing project
Unpublished (2000)
REFERENCE 2 (bases 1 to 1229)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao
Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomicshri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; CDNA full insert
sequencing: Research Association for Biotechnology; CDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.

FEATURES
Location/Qualifiers
source 1..1229
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1002458"
/clone_lib="HEMBA1"
/dev_stage="embryo, 10 weeks"
/tissue_type="whole embryo, mainly head"
/note="cloning vector: pME18SFL3"
191..553
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91443.1"
/db_xref="GI:7021946"

BASE COUNT 351 a 243 c 296 g 339 t
ORIGIN
KPGNIRGVRRLVPRVWPAIRNRIEHNARDDVRFVGOMKEIKRTRSQOMRHY
MRQTPEDNDHDFCLIP"

alignment_scores:
Quality: 355.50 Length: 120
Ratio: 3.665 Gaps: 3
Percent Similarity: 80.833 Percent Identity: 59.167

alignment_block:

US-09-327-750D-35 x AK000959

Align seg 1/1 to: AK000959 from: 1 to: 1229

```

1 MetAlaSerLysLysGlnValIleLeu...AspLeuThrValGluLysAs 16
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 ATGGAGTCCAAAGAGGAAGTACGGCAACAATCTCAACGGGGAATGC 240
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
16 physLysAspLysArgGlyGly...LysAlaSerLysGlnSerGluGlu 32
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 CCAACAAGAAACCAAGAGGGGAGGAGCCGCCAGCAGAAATGAAGA 290
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
32 LuProHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 48
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
291 AATCCCGCCATTTGGGAGGGGGTGAAGCCAGAGCCTGGAGGAATATC 340
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 64
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
341 AGCGGGGGCGAGTGTAGCGACTTGTCCCTAATTTTCGATGGCCATACC 390
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 OAsnArgHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheV 81
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 TAATAGGCATATTGAGCACAAATGAAGCGAGAGATGATGTAGAAGGTTT 440
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 alValGlnGlyThrGluValLysArgLysThrThrGluGlnGlnValArg 97
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 TAGGCAGATGATGCAATCAAGAAAGACTAGGGAACAGCAGATGAGG 490
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 ProTyArgArgPheArgThrProGluProAspAsnHisTyArgPheCy 114
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 CACTATATGCGCTTCCAAACTCTGAACTGACACCACTATGACTTTTG 540
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 sleuilePro 117
  ||||| |||||
541 CUTCATACCT 550

```

seq_name: gb_pat:AX100231

seq_documentation_block:

LOCUS AX100231 1364 bp DNA PAT 02-APR-2001

DEFINITION Sequence 32 from Patent WO0119860.

ACCESSION AX100231

VERSION AX100231.1 GI:13539115

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1364)

Tang, F.T., Hillman, J.L., Yue, H., Reddy, R., Lal, P., Shah, P.,

Azimzai, Y., Baughn, M.R., Lu, D.A., Bandman, O., Shih, L.L. and

Patterson, C.

TITLE Proteins associated with cell differentiation

JOURNAL Patent: WO 0119860-A 32 22-MAR-2001;

Incite Genomics, Inc. (US)

FEATURES

source

1..1364

/organism="Homo sapiens"

/db_xref="taxon:9606"

/note="Incite ID No: 1990956CB1"

BASE COUNT 411 a 269 c 322 g 362 t

ORIGIN

alignment_scores:

Quality: 355.50 Length: 120

Ratio: 3.665 Gaps: 3

Percent Similarity: 80.833 Percent Identity: 59.167

alignment_block:

US-09-327-750D-35 x AX100231

Align seg 1/1 to: AX100231 from: 1 to: 1364

1 MetAlaSerLysLysGlnValIleLeu...AspLeuThrValGluLysAs 16

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

250 ATGGAGTCCAAAGAGGAAGTACGGCAACAATCTCAACGGGGAATGC 299

```

16 physLysAspLysArgGlyGly...LysAlaSerLysGlnSerGluGlu 32
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CCAACAAGAAACCAAGAGGGGAGGAGCCGCCAGCAGAAATGAAGAAG 349
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
32 LuProHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 48
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 AATCCCGCCATTTGGGAGGGGGTGAAGCCAGAGCCTGGAGGAATATC 399
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 64
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 AGCGGGGGCGAGTGTAGCGACTTGTCCCTAATTTTCGATGGCCATACC 449
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 OAsnArgHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheV 81
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 TAATAGGCATATTGAGCACAAATGAAGCGAGAGATGATGTAGAAGGTTT 499
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 alValGlnGlyThrGluValLysArgLysThrThrGluGlnGlnValArg 97
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 TAGGCAGATGATGCAATCAAGAAAGACTAGGGAACAGCAGATGAGG 549
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 ProTyArgArgPheArgThrProGluProAspAsnHisTyArgPheCy 114
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
550 CACTATATGCGCTTCCAAACTCTGAACTGACACCACTATGACTTTTG 599
  ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 sleuilePro 117
  ||||| |||||
600 CUTCATACCT 609

```

seq_name: gb_pr:HS635G19

seq_documentation_block:

LOCUS HS635G19 69548 bp DNA PRI 23-NOV-1999

DEFINITION Human DNA sequence from clone 635G19 on chromosome Xq22.1-22.3

Contains a LAMR1 (Laminin Receptor 1 (67kD) (RPSA, 40S Ribosomal

Protein SA, P40)) pseudogene and part of a novel protein. Contains

ESTs and GSSs, complete sequence.

AL035494

ACCESSION AL035494.8 GI:4775622

VERSION HTG; 40S Ribosomal Protein P40; 40S Ribosomal Protein SA; Cpg

KEYWORDS Island; Laminin Receptor 1; LAMR1; RPSA.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 69548)

AUTHORS Bird, C.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On May 11, 1999 this sequence version replaced gi:4678749.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

This sequence was generated from part of bacterial clone contigs of

human chromosome X, constructed by the Sanger Centre Chromosome X

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/ChrX

635G19 is from the library RPC14 constructed at the Roswell Park

Cancer Institute by the group of Pieter de Jong. For further

details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>
 IMPORTANT: This sequence is not the entire insert of clone 635G19.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring submissions.
 The true left end of clone U65A4 (281014) is at 69549 in this
 sequence. The true right end of clone U101D3 (285997) is at 100 in
 this sequence.

FEATURES

Location/Qualifiers

source

1..69648
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="q22.1-22.3"
 /clone="RP4-635G19"
 /clone.lib="RPC1-4"

repeat_region
 439..483
 /note="TAR1 repeat: matches 707..749 of consensus"
 501..561
 /note="AluJb repeat: matches 137..311 of consensus"
 662..897
 /note="118 copies 2 mer aa 70% conserved"
 915..1331
 /note="MSTA repeat: matches 1..419 of consensus"
 1334..1517
 /note="92 copies 2 mer ga 77% conserved"
 1927..2020
 /note="MIR repeat: matches 109..217 of consensus"
 2066..2262
 /note="L1ME2 repeat: matches 5825..6022 of consensus"
 2263..2574
 /note="AluSp repeat: matches 1..313 of consensus"
 2575..2819
 /note="L1ME2 repeat: matches 5583..5825 of consensus"
 2821..3039
 /note="Alu repeat: matches 1..311 of consensus"
 3104..3384
 /note="L1ME2 repeat: matches 5304..5591 of consensus"
 3391..4007
 /note="L1MB1 repeat: matches 5558..6162 of consensus"
 4012..4137
 /note="Tigger2a repeat: matches 11..123 of consensus"
 4141..4252
 /note="L1 repeat: matches 2920..3031 of consensus"
 4265..4350
 /note="HY4 repeat: matches 1..87 of consensus"
 4350..4678
 /note="MER74A repeat: matches 28..369 of consensus"
 4914..5095
 /note="MER5A repeat: matches 8..189 of consensus"
 5224..5555
 /note="L2 repeat: matches 2423..2746 of consensus"
 6040..6086
 /note="L2 repeat: matches 1676..1720 of consensus"
 7413..7478
 /note="Alu repeat: matches 239..304 of consensus"
 7628..7928
 /note="AluJb8 repeat: matches 1..316 of consensus"
 7952..7987
 /note="18 copies 2 mer aa 81% conserved"
 7990..8305
 /note="AluJb repeat: matches 1..309 of consensus"
 8436..8538
 /note="MIR repeat: matches 99..212 of consensus"
 10836..11442
 /note="L1MD repeat: matches 1357..1972 of consensus"
 11443..11571
 /note="FLAN4 repeat: matches 2..129 of consensus"
 11572..11946
 /note="L1MD repeat: matches 977..1357 of consensus"
 11947..12245
 /note="Alu repeat: matches 1..301 of consensus"
 12246..13098

repeat_region
 13094..13223
 /note="L1MD repeat: matches 275..977 of consensus"
 13523..13820
 /note="L1MD repeat: matches -17..113 of consensus"
 13823..13932
 /note="MLTID repeat: matches 177..505 of consensus"
 13936..13977
 /note="55 copies 2 mer ct 77% conserved"
 13978..14240
 /note="MLTID repeat: matches 201..241 of consensus"
 14245..14534
 /note="AluSg1 repeat: matches 1..270 of consensus"
 14535..14721
 /note="AluSc repeat: matches 1..290 of consensus"
 14724..15604
 /note="MLTID repeat: matches 2..201 of consensus"
 15605..15909
 /note="L1MC/D repeat: matches 4646..5565 of consensus"
 16069..16363
 /note="AluSx repeat: matches 24..312 of consensus"
 16482..16641
 /note="AluSx repeat: matches 1..291 of consensus"
 16671..16953
 /note="AluJb repeat: matches 136..295 of consensus"
 17184..17280
 /note="AluSx repeat: matches 1..284 of consensus"
 17326..17377
 /note="L1P repeat: matches 4499..4619 of consensus"
 17704..17863
 /note="26 copies 2 mer at 79% conserved"
 17864..18175
 /note="AluSc repeat: matches 118..295 of consensus"
 18176..18298
 /note="AluJb8 repeat: matches 1..318 of consensus"
 18848..18873
 /note="AluSc repeat: matches 1..118 of consensus"
 18876..19171
 /note="113 copies 2 mer tg 100% conserved"
 19183..19232
 /note="AluSg repeat: matches 1..295 of consensus"
 19246..19533
 /note="25 copies 2 mer ac 100% conserved"
 19539..19837
 /note="AluY repeat: matches 5..292 of consensus"
 20016..20194
 /note="MER74B repeat: matches 61..368 of consensus"
 20394..20692
 /note="MER74A repeat: matches 295..487 of consensus"
 20701..20809
 /note="AluSg repeat: matches 1..302 of consensus"
 20877..21061
 /note="LTR1 repeat: matches 682..785 of consensus"
 21186..21636
 /note="LTR1 repeat: matches 1..195 of consensus"
 21609..21737
 /note="LTR29 repeat: matches 3..498 of consensus"
 21771..22076
 /note="MER34 repeat: matches 409..543 of consensus"
 22878..22974
 /note="AluY repeat: matches 2..308 of consensus"
 23452..23582
 /note="MIR repeat: matches 92..191 of consensus"
 23583..23624
 /note="FLAN_C repeat: matches 1..131 of consensus"
 23735..24041
 /note="21 copies 2 mer ta 76% conserved"
 24045..24072
 /note="AluSg1 repeat: matches 1..308 of consensus"
 25274..25582
 /note="14 copies 2 mer tt 89% conserved"
 25274..25582
 /note="AluJb repeat: matches 1..289 of consensus"
 25583..25747
 /note="match: GSS B63622"

misc_feature
 12246..13098


```

|||||.....|
316 CGCCGACTGCCCTAATTTTCATGGGCCATACCAATAGGAGATCAA 365
69 pArgAsnGluGly.....GlyGluAspValGlyArgPheValValGlnG 84
   |||||.....|
366 TGATGGGATGGGTGGAGATGGAGATGATGGAATATTCATGGAGAGA 415
84 lyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrArg 100
   |||||.....|
416 TGAGAGAAATCAGAGAAACTTAGGAGCTGCAGTTGAGGAATTGCTG 465
101 ArgPheArgThrProGluProAsnHisTyrAsp.....Ph 113
   |||||.....|
466 CGTATCCTTATGGGAGGCTCTTAATCACCATGACCATGATGAATT 515
113 eCysLeuIlePro 117
   |||||.....|
516 TTGCCTTATGCCT 528

```

seq_name: gb_pr:HUMOGC

```

seq_documentation_block:
LOCUS HUMOGC 891 bp mRNA PRI 07-MAR-1995
DEFINITION Human unknown protein from clone PHGR74 mRNA, complete cds.
ACCESSION M38188 X56942
VERSION M38188.1 GI:189378
KEYWORDS Human ovarian granulosa cell line, cDNA to mRNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 891)
AUTHORS Rapp, G., Freudenstein, J., Klaudiny, J., Mucha, J., Wempe, F., Zimmer, M., and Scheit, K.H.
TITLE Characterization of three abundant mRNAs from human ovarian granulosa cells
JOURNAL DNA Cell Biol. 9 (7), 479-485 (1990)
MEDLINE 91025550
COMMENT Draft entry and computer-readable sequence for [DNA 9, 479-485 (1990)] kindly submitted by K.H.Scheit, 27-AUG-1990.
FEATURES
source
1..891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PHGR74"
/cell_type="granulosa"
/tissue_type="ovary"
<1..891
/note="protein of unknown function"
312..547
/note="protein of unknown function"
/codon_start=1
/protein_id="AAA63232.1"
/db_xref="GI:189379"
/translations="MANIHOENEEMEPOMGEEDRPLGGEGHOPAGNRGQARRLA
PNFEWATPNQINDGMDGDDMEIPMEERIRKRLRELQLRNLRLMGLSELNHD
HHDEFCLMP"
361..534
/note="protein of unknown function"
/codon_start=1
/protein_id="AAA63233.1"
/db_xref="GI:189380"
/translations="MERKTALEEVKATSLQIDGDLADLPITFDGYPPIGRSMGW
VEMINKYSWRR"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN

```

alignment_scores:
 Quality: 175.00 Length: 121
 Ratio: 2.273 Gaps: 4
 Percent Similarity: 63.636 Percent Identity: 36.364

alignment_block:

US-09-327-750D-35 x HUMOGC

Align seg 1/1 to: HUMOGC from: 1 to: 891

```

4 LysLysGlnValIleLeu...AspLeuThrValGluLysAspLysLysAs 19
   |||||.....|
297 AAAAAAATCTCATCTGCAAAATATTCACCAAGAAACGAAGAGATGGA 346
19 pLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisLisL 36
   |||||.....|
347 GCAG.....CCTATGCAGAAATGGAGAGAGACCCGCTT 381
36 euGluGluValGluAsnLysLysProGlyGlyAsnValArgLysVal 52
   |||||.....|
382 TGGAGGAGGTGAAGCCACCGCTGCAGAAATCGACGGGACAGGCT 431
53 ArgArgLeuValProAsnPheLeuTTPAlaIleProAsnArgHisValas 69
   |||||.....|
432 CGCCGACTTCCCTAATTTTCATGGGCCATACCAATAGGAGATCAA 481
69 pArgAsnGluGly.....GlyGluAspValGlyArgPheValValGlnG 84
   |||||.....|
482 TGATGGGATGGGTGGAGATGGAGATGATGGAATATTCATGGAGAGA 531
84 lyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrArg 100
   |||||.....|
532 TGAGAGAAATCAGAGAAACTTAGGAGGCTGCAGTTGAGGAATTGCTG 581
101 ArgPheArgThrProGluProAsnHisTyrAsp.....Ph 113
   |||||.....|
582 CGTATCCTTATGGGAGGCTCTTAATCACCATGACCATGATGAATT 631
113 eCysLeuIlePro 117
   |||||.....|
632 TTGCCTTATGCCT 644

```

seq_name: gb_pr:AF187064

seq_documentation_block:

LOCUS AF187064 891 bp mRNA PRI 11-JUN-2000

DEFINITION Homo sapiens p75NTR-associated cell death executor (NADE) mRNA,

complete cds.

ACCESSION AF187064

VERSION AF187064.1 GI:8452893

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homi

REFERENCE 1 (bases 1 to 891)

AUTHORS Mukai, J., Hachiya, T., Shoji, Hoshino, S., Kimura, M. T., Nadano, D.,

Suvanto, P., Hanaoka, T., Li, Y., Irie, S., Greene, L. A. and Sato, T. A.

NADE, a p75NTR-associated cell death executor, is involved in

signal transduction mediated by the common neurotrophin receptor

p75NTR

J. Biol. Chem. 275 (23), 17566-17570 (2000)

MEDLINE 20298829

REFERENCE 2 (bases 1 to 891)

AUTHORS Mukai, J., Hachiya, T., Hoshino, S., Kimura, M., Nadano, D.,

Hanaoka, T., Li, Y., Irie, S. and Sato, T.

Direct Submission

TITLE Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia

University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA

JOURNAL Location/Qualifiers

FEATURES

source

1..891

/organism="Homo sapiens"

/db_xref="taxon:9606"

1..891

/gene="NADE"

312..647

/gene="NADE"


```

alignment_block:
US-09-327-750D-35 x HSV351F8  ..
Align seg 1/1 to: HSV351F8 from: 1 to: 45678

11 LeuthrValGluLysAspLysLysAspLysArgGlyGlyLysAlaSerIly 27
19408 TTATCTTCTAGGAAGAAAAAAGAAAAACAGGAAAAAACTCAACAT 19457
27 scInserGluGluProHisHisLeuGluGlu.....ValGluA 41
19458 GGAATATGTCCTCCCAAGAAAAACAAGTTGTGGAGAGGCCCTGCAGA 19507
41 sn.....LysLysProGlyGlyAsn 47
19508 ATGAAGCCCGCGCTTTAGGAGGTGGTGAATACAGGAGCGCTGGAGGAAT 19557
48 VallArgArgLysValArgGluValProAsnPhaLeuTirAlaIlePr 64
19558 GTTAAAGGGGTTGGGCTCCACCTGCCCGGGTTTGGAGAGGATGTGCC 19607
64 oAsnArgHisValAspArg.....AsnGluGlyGlyGluAspV 77
19608 CAATAGGCTTCGATAACATTCATATAGATAGATGAGATGGAGATGATA 19657
77 alGlyArgPheValValGluGlyThrCluValLysArgLysThrThrGlu 93
19658 TGAACGCTTCATGAGGATGATGAGAGCTAAGGAGGAAATTAGGNA 19707
94 GlnGlnValArgProTyrArgPheArgThrProGluPro..... 107
19708 CTTAGTTGAGTACAGTCTGCGCATCTTATAGGGGACCTCCTCACCA 19757
108 AsnAsnHisTyrAspPheCysLeuIlePro 117
19758 TGATCATGATGAGTTTGCCTTATGCCT 19788

seq_name: gb_sts:G35294

seq_documentation_block: 477 bp DNA STS 02-OCT-1997
LOCUS G35294
DEFINITION human STS SHGC-37409, sequence tagged site.
ACCESSION G35294
VERSION G35294.1 GI:2459462
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 477)
Myers, R.M.
Unpublished (1997)
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu
Primer A: AACATCTTTCCATGAAGATGATG
Primer B: CTTTGGCATCTTCTGCAA
STS size: 106
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 15 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from N51315
-- Washington University/Merck EST sequence.
FEATURES
source
1..477
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
44..149
primer_bind'
44..67
primer_bind
complement(130..149)
BASE COUNT 153 a 115 c 79 g 130 t
ORIGIN

alignment_scores:
Quality: 120.50 Length: 83
Ratio: 2.410 Gaps: 2
Percent Similarity: 60.241 Percent Identity: 36.145

alignment_block:
US-09-327-750D-35 x G35294/rev ..
Align seg 1/1 to reverse of: G35294 from: 1 to: 477

42 LysLysProGlyGlyAsnValArgArgLysValArgGluValProAs 58
474 CAGGAGCGCTGGAGAAATGTTAAAGGGGTTTGGGCTCCACCTGCCCGGG 425
58 nPheLeuTirAlaIleProAsnArgHisValAspArg.....A 71
424 TTTTGAGAGAGGATGTCGCAATAGGCTTTCGATAACATTCATGATAG 375
71 snGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluVal 87
374 ATGGACATGAGATGATATGGACGTTTCATGGAGAGATGAGAGACTA 325
88 LysArgLysThrThrGluGlnValArgProTyrArgArgPheArgTh 104
324 AGGAGAAATAGGGAACCTTCAGTTGAGGTACAGTCTGCGCATTCCTAT 275
104 rProGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 117
274 AGGGAGCGCTCCTCACCATGATCATGATGAGTTTGCCTTATGCCT 226

seq_name: gb_pr:HS714B7

seq_documentation_block:
LOCUS HS714B7 98274 bp DNA PRI 12-DEC-1999
DEFINITION Human DNA sequence from clone CTA-714B7 on chromosome 22q12.2-13.2
Contains pseudogene similar to part of COX7B (cytochrome c oxidase
subunit VIIb), a novel pseudogene, ESTs, STS and GSSs, complete
sequence.
ACCESSION Z99755
VERSION Z99755.1 GI:3036782
KEYWORDS HTG; COX7B; cytochrome c oxidase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 98274)
Connor, R.
Direct Submission

```

JOURNAL

Submitted (08-DEC-1999), Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk. Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Apr 8, 1998 this sequence version replaced gl:2578134. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/Celegans/wormpep> CTA-714B7 is from the human BAC library described in U-J. Kim et al. (1996) Genomics 34, 213-218.

VECTOR: pBel0BAC11

This sequence is the entire insert of clone CTA-714B7.

FEATURES

source

Location/Qualifiers
1..98274
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="22"
/map="q12.2-13.2"
/clone="CTA-714B7"
/clone_lib="CIR978SK-A2"
109..136
/note="14 copies 2 mer ac 100 conserved"

repeat_region

repeat_region 183..258
/note="L1MC5 repeat: matches 7840..7913 of consensus"

misc_feature

1177..1446
/note="match: GSS: Em:AQ373724"

repeat_region

1301..1356
/note="L2 repeat: matches 2693..2750 of consensus"

repeat_region

2568..5038
/note="L1MD1 repeat: matches 3673..6170 of consensus"

repeat_region

5036..5486
/note="L1MEC repeat: matches 2052..2207 of consensus"

repeat_region

5496..6885
/note="L1MD1 repeat: matches 1..1538 of consensus"

repeat_region

6896..7185
/note="L2 repeat: matches 2356..2710 of consensus"

misc_feature

complement(7187..7611)
/note="match: GSS: Em:AQ566001"

repeat_region

7369..7594
/note="MIR repeat: matches 34..256 of consensus"

misc_feature

complement(7713..8092)
/note="match: GSS: Em:AQ071449"

repeat_region

7904..7978
/note="MIR repeat: matches 20..95 of consensus"

repeat_region

8086..8300
/note="MIR repeat: matches 8..255 of consensus"

misc_feature

8154..8565
/note="match: GSS: Em:AQ059599"

repeat_region

8349..8864
/note="AluX repeat: matches 1..280 of consensus"

misc_feature

complement(8897..9290)
/note="match: GSS: Em:AQ343214"

repeat_region

9235..9361
/note="L2 repeat: matches 2619..2749 of consensus"

misc_feature

9378..9890
/note="match: GSS: Em:AQ284832"

repeat_region

10135..10428
/note="AluX repeat: matches 1..294 of consensus"

repeat_region

10429..10563
/note="AluSq/x repeat: matches 1..135 of consensus"

repeat_region

10811..10868
/note="L2 repeat: matches 2648..2705 of consensus"

repeat_region

10819..11064
/note="MIR repeat: matches 2..255 of consensus"

repeat_region

11112..11124
/note="MIR repeat: matches 11..255 of consensus"

repeat_region

11869..12012
/note="MIR repeat: matches 20..167 of consensus"

repeat_region

12925..13223
/note="AluY repeat: matches 1..300 of consensus"

repeat_region

13290..13495
/note="MER58A repeat: matches 9..214 of consensus"

misc_feature

13338..13833
/note="match: GSS: Em:AQ568864"

repeat_region

13688..13749
/note="MIR repeat: matches 77..139 of consensus"

repeat_region

13840..13966
/note="MIR repeat: matches 22..147 of consensus"

repeat_region

14162..14245
/note="MIR repeat: matches 46..136 of consensus"

repeat_region

14925..15058
/note="MIR repeat: matches 48..191 of consensus"

repeat_region

15994..16040
/note="MIR repeat: matches 102..144 of consensus"

repeat_region

16116..16184
/note="MER5B repeat: matches 3..70 of consensus"

repeat_region

16897..17005
/note="MIR repeat: matches 9..118 of consensus"

repeat_region

17618..17700
/note="MIR repeat: matches 108..192 of consensus"

repeat_region

17748..17912
/note="MIR repeat: matches 73..245 of consensus"

repeat_region

18657..18871
/note="MIR repeat: matches 35..252 of consensus"

repeat_region

18927..19230
/note="AluX repeat: matches 1..301 of consensus"

repeat_region

19492..19559
/note="MIR repeat: matches 2..67 of consensus"

repeat_region

19560..19821
/note="AluY repeat: matches 34..295 of consensus"

repeat_region

19822..19959
/note="MIR repeat: matches 67..212 of consensus"

repeat_region

19987..20286
/note="AluY repeat: matches 1..300 of consensus"

repeat_region

20316..20512
/note="MIR repeat: matches 82..262 of consensus"

repeat_region

20513..20741
/note="L1MD1 repeat: matches 5970..6224 of consensus"

repeat_region

20769..21074
/note="AluSq1 repeat: matches 1..303 of consensus"

repeat_region

21085..21207
/note="L1M2 repeat: matches 1776..1898 of consensus"

repeat_region

21212..21391
/note="L1MD1 repeat: matches -6..178 of consensus"

repeat_region

21392..21471
/note="MIR repeat: matches 3..86 of consensus"

repeat_region

22146..22441
/note="AluX repeat: matches 1..296 of consensus"

repeat_region

22526..22610
/note="L1R16C repeat: matches 267..349 of consensus"

repeat_region

22940..23204
/note="HERV16 repeat: matches 378..650 of consensus"

repeat_region

23205..23514
/note="AluSq repeat: matches 1..311 of consensus"

repeat_region

23919..24194
/note="AluX repeat: matches 36..311 of consensus"

repeat_region

24496..24590
/note="L1R16C repeat: matches 253..349 of consensus"

repeat_region

24871..24916


```

452 AGACGTACAGAGTTTGGGGTGATGTGAGACAGCTCATGGAGAGCTG 501
92 ThrGluGlnValArgProTyrArgArgPheArgThrProGluPro.. 107
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
502 AGGAAAGGAGCTGAGCCACACAGCTGGGGGGTTCAGCTAGTACCGGCC 551
108 .....AspAsnHisTyrAspPheCysLeuLeuPro 117
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
552 TCATCATGACCCACCATGATGAGTTTGGCTCATGCC 588
seq_name: gb_ro:AF051347

seq_documentation_block:
LOCUS AF051347 838 bp mRNA ROD 01-OCT-1998
DEFINITION Mus musculus REX-3 mRNA, complete cds.
ACCESSION AF051347
VERSION AF051347.1 GI:3510642
KEYWORDS
SOURCE house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 838)
Faria,T.N., LaRosa,G.J., Wilen,E., Liao,J. and Gudas,L.J.
Characterization of genes which exhibit reduced expression during
the retinoic acid-induced differentiation of F9 teratocarcinoma
cells: involvement of cyclin D3 in RA-mediated growth arrest
Mol. Cell. Endocrinol. 143 (1-2), 155-166 (1998)
99021197
JOURNAL MEDLINE
REFERENCE 2 (bases 1 to 838)
AUTHORS Faria,T.N., LaRosa,G., Wilen,E., Liao,L. and Gudas,L.J.
TITLE Direct Submission
JOURNAL Submitted (26-FEB-1998) Pharmacology, Cornell University Medical
College, 1300 York Avenue, New York, NY 10021, USA
FEATURES
source
1..838
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="F9 teratocarcinoma"
229..579
/note="retinoic acid reduced gene expression"
/codon_start=1
/product="REX-3"
/protein_id="AAC61929.1"
/db_xref="GI:3510643"
/translation="MENDHOKKEEKPKQDITRREPAVALISEAGKNCAPRGRRRRF
RVROPIAHYRWDLQMQRVGEQGRMRRENVQRFEGDVRQLMEXLRERQLSHSLRAVSTD
PPHDDHDEFCLRP"
BASE COUNT 238 a 181 c 251 g 168 t
ORIGIN

alignment_scores:
Quality: 107.50 Length: 129
Ratio: 1.473 Gaps: 6
Percent Similarity: 56.589 Percent Identity: 30.233

alignment_block:
US-09-327-750D-35 x AF051347
..
Align seg 1/1 to: AF051347 from: 1 to: 838
1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAsp.. 16
||| ||||| |||::: |||::: |||::: |||::: |||::: |||::: |||
193 ATGAGTCCAAGATCAAGCGGTGAAAAATCTCAACATGGAGATGACCA 242
17 ....LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGlu 32
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
243 TCAGAAAAAGGAGGAGAGGAGAAAGCCACAGATACCATCAGAGGG 292
32 luProHisHisLeu.....GluGluValGluAsnLysLysPro... 44
|||||
293 AGCCAGCTGTGGCCCTGATCTCCGAGAGTGGCAAAAACACTGTGGCCCTAGA 342

```

```

45 GlyClyAsnValArgArgLysValArgArgLeuValProAsnPheLeuTr 61
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
343 GGAGGTCCGAGCGGTTCCGGGTTCCGGAGCCCATCGCTACTATAGATG 392
61 pAlaIleProAsnArgHisValAspArgAsnGluGly.....GlyG 75
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
393 GGACCTGATGCAGAGG...GTTGGGAGCCCCCAGGGAAGGATGAGAGAGG 439
75 luAspValGlyArgPheValValGlnGlyThrGluValLysArgLysThr 91
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
440 AGAACGTACAGAGTTTGGGGTGATGTGAGACAGCTCATGGAGAGCTG 489
92 ThrGluGlnGlnValArgProTyrArgArgPheArgThrProGluPro... 107
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
490 AGGAAAGGTCAGCTGAGCCACACAGCTGCGGGGTTAGCAGCTGACCCGCC 539
108 .....AspAsnHisTyrAspPheCysLeuLeuPro 117
|||::|||::: |||::: |||::: |||::: |||::: |||::: |||
540 TCATCATGACCCACCATGATGAGTTTGGCTCATGCC 576

```

```
34 erHisHisLeuGluGluValGluAlaLysArgHisSerLysProGlyGlyAsnValArg 50
|||||
268 CCCACCATCTGGAGAGGTTGAAACAGAGGCTGGGGGAATATGTCGGA 317
|||||
51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
|||||
318 AGAAGTC. AGGGACTTGTGCTAACTTCTCTGGGCCATACCAATAG 366
|||||
67 gHisValAspHisSerCluclyGlyGluGluValGlyArgPheValClyG 84
|||||
367 GCATGTTGATCGCAATGAGGGGAGGAGGATGTTGGGAGATTGCTAGTGC 416
|||||
84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgPro 99
|||||
417 AGGACACAGAGTCACAGAGAAAGACTACGAGCAGCAGGTCAGGCCT 463
|||||
seq_name: gb_est1.AL118340
```

```
seq_documentation_block:
LOCUS AL118340 422 bp mRNA EST 23-SEP-1999
DEFINITION V9112b26 Beddington mouse dissected endoderm Mus musculus cDNA
clone 528_13N20 5', mRNA sequence.
ACCESSION AL118340
VERSION AL118340.1 GI:5920179
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 422)
Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and
Beddington,R.S.
Isolation of novel tissue-specific genes from cDNA libraries
representing the individual tissue constituents of the gastrulating
mouse embryo
Development 121 (8), 2479-2489 (1995)
95401865
Contact: Ruiz,P., Lehrach,H. and Avner,P.
EEC Mouse Transcript Mapping Consortium
Genoscope - CNS
2, rue Gaston Cremieux, CP 5706, 91057 Evry CEDEX, France
clone available from Ressourcenzentrum, Heubnerweg 6, D-14059
Berlin, Germany. Web site http://www.rzpd.de
Seq primer: CCGTCCGGAATTCCTCCGGT.
Location/Qualifiers
1..422
/organism="Mus musculus"
/strain="C57Bl6 x DBA"
/db_xref="taxon:10090"
/clone="528_13N20"
/clone_lib="Beddington mouse dissected endoderm"
/tissue_type="dissected endoderm"
/dev_stages="7.5 dpc"
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Cloned
unidirectionally. ~ High quality sequence only submitted. ~
Average insert size: 1.2 kb (range: 0.2 - 2.kb)"
```

```
BASE COUNT 128 a 86 c 143 g 65 t
ORIGIN
```

```
alignment_scores:
Quality: 407.00 Length: 90
Ratio: 4.788 Gaps: 0
Percent Similarity: 94.444 Percent Identity: 87.778
```

```
alignment_block:
US-09-327-750D-34 x AL118340
Align seg 1/1 to: AL118340 from: 1 to: 422
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
```

```
|||||
146 ATGGCATCCAAATTTAAACAAGTCATACTGGATCTCACTGTTGGAGAAAGA 195
|||||
17 pLysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGluS 34
|||||
196 CAAAAAAGACAAAGAGAGGTGGGAAGGCTCCAAACAAAGTGAAGAGAAC 245
|||||
34 erHisHisLeuGluGluValGluAlaLysArgHisSerLysProGlyGlyAsnValArg 50
|||||
246 CCCACCATCTGGAGAGGTTGAAACAGAGGCTGGGGGAATATGTCGGA 295
|||||
51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
|||||
296 AGGAAAGTCAGGCGACTTGTGCTAACTTCTCTGGGCCATACCAATAG 345
|||||
67 gHisValAspHisSerCluclyGlyGluGluValGlyArgPheValClyG 84
|||||
346 GCATGTTGATCGCAATGAGGGGAGAGGATGTTGGGAGATTGCTGCTGC 395
|||||
84 InValMetGluAlaLysArg 90
|||||
396 AGGACACAGAGTCACAGAGA 415
```


508 AGGCGTTTCCGAACCCCGGAACCTCACAATCATTACGACTTTTGCTCAT 557

117 ePro 118
|||||
558 ACCT 561

seq_name: gb_est2.BG083261

seq_documentation_block: 792 bp. mRNA EST 26-JAN-2001
LOCUS BG083261 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
DEFINITION H3086C08-5, mRNA sequence.
ACCESSION BG083261
VERSION BG083261.1 GI:12565829
KEYWORDS EST.
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 792)
AUTHORS Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
T.S., Carter, M.G. and KO, M.S.H.
TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other_ESTs: H3086C08-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdnaelgusn.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3086 Row: C Column: 08
Seq primer: -21M13 Reverse
High quality sequence stop: 792
POLYA-No.

FEATURES

source Location/Qualifiers
1. .792
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3086C08"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
BASE COUNT 247 a 153 c 210 g 179 t 3 others
ORIGIN

alignment_scores:

Quality: 546.00 Length: 118
Ratio: 4.964 Gaps: 0

Percent Similarity: 93.220 Percent Identity: 87.288

alignment_block:

US-09-327-750D-34 x BG083261 ..

Align seg 1/1 to: BG083261 from: 1 to: 792

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
|||||
190 ATGCATCCCAAAATTAACAAGTCACTACGATCTCACTGTGGAGAAGA 239
|||||
17 pLysLysAsnLysGlyLysAlaSerLysGlnSerGluGluGlu 34
|||||
240 CAAAAAGACAAAAAGGTGGCAAGGCTCCAAACAAGTGAAGAAGAAC 289
|||||
34 erHisLysLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
|||||
290 CCCACCATCTCGAAGAGGTTCAAAACAAGAGCCCTGGGGAAATGTCGA 339
|||||
51 ArgLysValArgArgLeuValProAsnPhelLeuTrpAlaLleProAsnAr 67
|||||
340 AGGAAAGTCAGCGGACTTGTGCTTAACCTTCTCTGGCCATACCAAAATAG 389
|||||
67 gHisValAspHisSerGluGlyGlyGluValGlyArgPheValGlyG 84
|||||
390 GCATGTTGATCGCAATGAAGGGGAGAGGATGTTGGGAGATTGTAGTGC 439
|||||
84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
|||||
440 AGGGACAGAGTCAAGAGAAGACTACGGAGCAGCAGCTGAGGCTTAC 489
|||||
101 ThrArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuIl 117
|||||
490 AGGCGTTTCCGAACCCCGGAACCTGACAATCATTACGACTTTTGCTCAT 539
|||||
117 ePro 118
|||||
540 ACCT 543

seq_name: gb_est1:AA473525

seq_documentation_block: 458 bp mRNA EST

LOCUS AA473525
DEFINITION v978409.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:872056 5',
similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD
PROTEIN HGR74. ; mRNA sequence.

ACCESSION AA473525

VERSION AA473525.1 GI:2201752

KEYWORDS EST.

SOURCE house mouse

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 458)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:511536

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 417.

Tur

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution Information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/resources.shtml
 Seq primer: -40UP from Gibco
 High quality sequence stop: 422.

FEATURES

source
 1. .612
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2849369"
 /clone_lib="NCI-CGAP_Lym12"
 /tissue_type="lymphoma, follicular mixed small and large
 cell"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
 Salt: Site_2: NotI; Cloned unidirectionally. Primer:
 Oligo dh. Average insert size 1.25 kb. Life Technologies
 catalog #: 11547-015"
 BASE COUNT 169 a 145 c 120 g 178 t
 ORIGIN

alignment_scores:
 Quality: 609.00 Length: 111
 Ratio: 5.486 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-327-750D-13 x AW512400/rev ..

Align seg 1/1 to reverse of: AW512400 from: 1 to: 612

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
 |||||
 581 ATGGCAAAATATTCACCGAGGAAACGAGAGATGGAGCGCTATGCAGAA 532
 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
 |||||
 531 TGGAGAGGAGAACCGCCCTTTGGGAGGAGGTGAAGGCCACCGCCCTGCAG 482
 34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAla 50
 |||||
 481 GAAATCGACGGGACAGGCTCGCGACTTGCCTTAATTTTCATGGGCC 432
 51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMe 67
 |||||
 431 ATACCAATAGGCAGATCAATGATGGATGGGTGGAGATGGAGATGATAT 382
 67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
 |||||
 381 GAAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTAGGGAGC 332
 84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
 |||||
 331 TGCAGTTGAGCAATTTGTCGCTATCCCTATGGGGAGCTCTCTAATCAC 282
 101 HisAspHisHisAspGluPheCysLeuMetPro 111
 |||||
 281 CATGACCATCATGATCAATTTTGCCTTATGCCT 249

seq_name: gb_estl:AI193112

seq_documentation_block:

LOCUS AI193112 615 bp mRNA EST 29-OCT-1998
 DEFINITION qe69g09.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
 IMAGE:1744288 3' similar to gb:M36188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AI193112
 VERSION AI193112.1 GI:3744321
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 845 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 445.

FEATURES

source
 1. .615
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1744288"
 /clone_lib="Soares_fetal_lung_NbHL19W"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGGCGCCCAATTTTCTTTT-3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Patima Ronaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NbHL19W."
 BASE COUNT 169 a 145 c 121 g 180 t
 ORIGIN

alignment_scores:
 Quality: 609.00 Length: 111
 Ratio: 5.486 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-327-750D-13 x AI193112/rev ..

Align seg 1/1 to reverse of: AI193112 from: 1 to: 615

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
 |||||
 584 ATGGCAAAATATTCACCGAGGAAACGAGAGATGGAGCGCTATGCAGAA 535
 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
 |||||
 534 TGGAGAGGAGAACCGCCCTTTGGGAGGAGGTGAAGGCCACCGCCCTGCAG 485
 34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAla 50
 |||||
 484 GAAATCGACGGGACAGGCTCGCGACTTGCCTTAATTTTCATGGGCC 435
 51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMe 67
 |||||
 434 ATACCAATAGGCAGATCAATGATGGATGGGTGGAGATGGAGATGATAT 385
 67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
 |||||
 384 GAAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTAGGGAGC 335
 84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
 |||||
 334 TGCAGTTGAGCAATTTGTCGCTATCCCTATGGGGAGCTCTCTAATCAC 285
 101 HisAspHisHisAspGluPheCysLeuMetPro 111

101 HisAspHisHisAspGluPheCysLeuMetPro 111

102(a)

human

```

|||||
284 CATGACCATCATGATGAATTTTGCCTTATGCCT 252
seq_name: gb_est1:AA576958

seq_documentation_block:
LOCUS AA576958 620 bp mRNA 12-SEP-1997
DEFINITION nm82404.s1 NCI-CGAP-Co9 Homo sapiens cDNA clone IMAGE:1074727 3'
similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
(HUMAN)), mRNA sequence.
ACCESSION AA576958
VERSION AA576958.1 GI:2354432
SOURCE EST.
ORGANISM human.
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 620)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 725 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 339.
Location/Qualifiers
1..620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1074727"
/clone_lib="NCI-CGAP-Co9"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo (Soares4
)."
```

BASE COUNT 169 a 144 c 116 g 191 t
ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750D-13 x AA576958/rev ..
Align seg 1/1 to reverse of: AA576958 from: 1 to: 620
1 MetalaasnlleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
|||||
602 ATGCCAAATATTCACAGGAAACGAAGAGATGGAGCGCTATGCGAGAA 553
17 nGlyGluGluAspArgProLeuGlyGlyGlyGluGlyHisGlnProAlaG 34
|||||
552 TGGAGAGAGAGACCGCCCTTTGGGAGGAGGTGAAGGCCACCAGCCTCGAG 503

```

34 lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
|||||
502 GAAATCGACGGGACAGGCTCGCCGACTTCCCTTAATTTTCGATGGGCC 453

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyGlyGlyGlyGlyGly 67
|||||
452 ATACCCCAATAGGAGATCATGATGGATGGTGGAGATGGAGATGATAT 403

67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
402 GGAATATTTCATGGAGAGATGAGAGAAATCAGAGAAAACCTTAGGGAGC 353

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
352 TGCAGTTGAGGAATGCTCTCGTATCCTTATGGGGAGAGCTCTCTAATCAC 303

101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
302 CATGACCATCATGATGAATTTTGCCTTATGCCT 270

seq_name: gb_est1:AI929106

seq_documentation_block:
LOCUS AI929106 662 bp mRNA 23-AUG-1999
DEFINITION au55b10.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2519611 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AI929106
VERSION AI929106.1 GI:5665070
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 662)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
High quality sequence stop: 437.
Location/Qualifiers
1..662
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2519611"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site 1: SstI; Site 2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGCTCAAGGATCTCTAATTAATTAATCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGCTCGAGTTTCTTTTCTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
```

FEATURES
source


```

/ad_xref="taxon:100930"
/clone_name="IMAGE:761628"
/clone_lib="barstead mouse pooled organs MPLR84"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH108"

```

17 nGlyGluGluAspArgProValGlyGlyGlyGluGlyHisGlnProAlaG 34
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
 225 TGGACAGGAAGACCGCCCTGTGGGAGGAGGTGAGGGCCACCAGCCTGCTG 274

51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||
325 GGCCAGGCTCGCCGACTTGCCCCTAACCTCCGATGGGCCATTCCCAACAG 374

84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
 |||||
 425 TGGAGGAGATGAGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 474

117 sAspGluPheCysLeuMetPro 124
|||||
525 TGATGAATTCTGCCTTATGCCT 546

seq_documentation_block:	LOCUS	AA272375	612 bp	mRNA	EST	26-MAR-1997
DEFINITION	v62g07.r1 Barstead mouse pooled organs MPRB4 MUS MUSCULUS CDNA					
	clone IMAGE:761628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL					
	13.0 KD PROTEIN HGR74 (HUMAN); mRNA sequence.					

KEYWORDS
SOURCE
EST.
house mouse

Mus-muscurosus	Eukaryota; Metazoa;	Chordata;	Carnivora; Vertebrata; Euteleostomi;
	Mammalia; Eutheria;	Rodentia;	Sciurognathi; Muridae; Mus.
1 (bases 1 to 612)			
Marra,M., Hillier,L., Allen,N., Bowles,M., Dietrich,N., Dubuque,T. Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Wyllie,E.			

COMMENT

wasnu-HHMI Mouse EST project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel. 314 286 1800

Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:462548

seq primer: -20113 rev2 EI FROM Allersham
High quality sequence stop: 507

TRADES	Location/Qualifiers
source	1 612

/organism="Mus musculus"

seq documentation block.

LOCUS
DEFINITION
W81757

clone IMAGE:

ACCESSION W81757

KEYWORDS EST.

SOURCE: house mouse.

11(mouse)

```
US-09-327-750D-12 x W46041
Align seg 1/1 to: W46041 from: 1 to: 492

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
13 ATGGCCAAATGTCACCAAGAAAGAGATGGAGCAGCCCTGCAGAA 62
17 nGlyGluGluAspArgProValGlyGlyGlyGluGlnProAlaG 34
63 TGACAGGAAGACCCCTGTGGAGGAGGTGAGGCGCCACCAAGCTG 112
34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArg 50
113 CAACACACACACACACACACACACACACACACACACACACAA 162
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIlePro 67
163 GCAGAGCTGCGCGAGCTTCCCTTCCGATGGCGCATTCACCAAG 212
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPhe 84
213 GCAGATGAATGACGGGTTGGGTGGAGATGGAGATGATGGAATG 262
84 etGluGluMetArgGluLeuArgArgLysLeuArgGluLeuArg 100
263 TGGAGGAGATGAGAGATGCGGAAAGCTTAGGAGCTACAGCTGAG 312
101 AsnCysLeuArgGluLeuMetGlyGluLeuSerAsnHisAsnHis 117
313 AATTGTCTACGATCTTATGGGAGCTGCTAACCAACCAAGATCAC 362
117 sAspGluPheCysLeuMetPro 124
363 TGATGAATTCCTGCTTATGCCT 384

seq_name: gb_est1:AI006575
```

```
seq_documentation_block:
LOCUS AI006575 538 bp mRNA EST 12-JUN-1998
DEFINITION ue15e06.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
IMAGE:1480450 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AI006575
VERSION AI006575.1 GI:3216184
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 538)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:928806
Seq primer: primer name ambiguous
High quality sequence stop: 433.
Location/Qualifiers
1..538
/organism="Mus musculus"
```

```
seq_documentation_block:
LOCUS BE334866 542 bp mRNA EST 14-JUL-2000
DEFINITION us90a11.y1 Soares mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:3325628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION BE334866
VERSION BE334866.1
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 538)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:928806
Seq primer: primer name ambiguous
High quality sequence stop: 433.
Location/Qualifiers
1..538
/organism="Mus musculus"
```

```
seq_name: gb_est1:BE334866
seq_documentation_block:
LOCUS BE334866 542 bp mRNA EST 14-JUL-2000
DEFINITION us90a11.y1 Soares mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:3325628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION BE334866
```

```
seq_documentation_block:
LOCUS BE334866 542 bp mRNA EST 14-JUL-2000
DEFINITION us90a11.y1 Soares mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:3325628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION BE334866
```

101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH117
 478 AATTGTCTACGATCCTTATGGGAGCTGTCTAACCAACAGATCACCA 527
 117 sAspGluPheCysLeuMetPro 124
 528 TGATGAATTCGCCCTATGCCT 549

seq_name: gb_est1.A1152323

seq_documentation_block:
 LOCUS A1152323 430 bp mRNA EST 30-SEP-1998
 DEFINITION u479a02.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
 IMAGE:1477034 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0

ACCESSION A1152323
 VERSION A1152323.1 GI:3680792
 KEYWORDS EST
 SOURCE house mouse
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 430)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:925390

Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 419.

FEATURES
 source

1. 430
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:1477034"
 /clone_lib="Soares_mammary_gland_NMLMG"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT7T3 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 135 a 111 c 113 g 71 t
 ORIGIN

alignment_scores:
 Quality: 681.00 Length: 124
 Ratio: 5.537 Gaps: 0
 Percent Similarity: 99.194 Percent Identity: 97.581

alignment_block:

US-09-327-7500-12 x A1152323

Align seg 1/1 to: A1152323 from: 1 to: 430

1 MetAlaAsnValHisGlnGlnAsnGluGluMetGluGlnProLeuGlnAs 17
 54 ATGGCAATGTCACCAAGAAACCAAGAGATGGAGCAGCCCTGCAGAA 103
 17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlnProAlaG 34
 104 TGGACAGGAAGACCGCCCTGTGGATGGAGGTGAGGCCACACAGCTGTG 153
 34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisHisArgArg 50
 154 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 203
 51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaIleProAsnAr 67
 204 GCCCAGGCTGCCGACCTTCCCTTAACCTCCGATGGCCCATTCACCAAG 253
 67 gGlnMetAsnAspGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 84
 254 GCAGATGAATGACGGTGGTGGAGATGGAGATGATATGGAATGTTCA 303
 84 etGluGluMetArgGluIleArgArgLysLeuArgGluGlnLeuArg 100
 304 TGGAGGAGATGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 353
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH117
 354 AATTGTCTACGATCCTTATGGGAGCTGTCTAACCAACAGATCACCA 403
 117 sAspGluPheCysLeuMetPro 124
 404 TGATGAATTCGCCCTATGCCT 425

seq_name: gb_est1.AW908751

seq_documentation_block:

LOCUS AW908751 503 bp mRNA EST 25-MAY-2000
 DEFINITION uf57a05.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
 IMAGE:1515440 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AW908751

VERSION AW908751.1 GI:8073984

KEYWORDS EST

SOURCE house mouse

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 503)

REFERENCE 1 (bases 1 to 503)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Other_ESTS: uf57a05.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:942292

Seq primer: -40RP from Gibco

High quality sequence stop: 468.

Location/Qualifiers

source

1. 503

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:1515440"

/clone_lib="Soares_mammary_gland_NMLMG"

/sex="female (lactating)"

/tissue_type="mammary gland"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not

Mar 12 09:01:15 2002
Cums 1, 2, 3, 4, 11 (mouse)¹³ 20, 21, 22, 23
102(a)

34	lyAsnAsnAsnAsnAsnAsnHIsAsnHIsAsnHIsAsnHIsArgArg	50
:	:::::	
277	CAACAAACAACAACAACAACAACAACAACAACAACAACAACA	326
51	GlyGlnAlaArgLeuAlaProAsnPhaArgTrpAlaIleProAsnAr	67
327	GGCCAGGCTCGCGGACTTGCGCCCTAACTTCGGATGGGCCATTCCCACAAG	376
67	gClnMetAsnAspGlyLeuGlyGlyAspGlyAspAspMetGluMetPheM	84
377	GCAGATGAATGACGGGTGGGTGGAGTGAGAGAATATGGAATAATTCCA	426
84	etCluGluMetArgGluIleArgLysLeuArgGluLeuGlnLeuArg	100
427	TGGAGGAGATCAGAGAGATCCGGAGAAAGCTTAGGGAGAGCTACAGCTGAGA	476
101	AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHIsHisAspHis	117
477	AATGTGCTACGGNMCCTTATGGGGAGCTGTCTAACCAACCACCATCACCA	526
117	sAspGluPheCysLeuMetPro	124
:	:::::	
527	TGATGAATTCGCTTATGCCT	548

```
seq_name: gb_ro:AF097440
seq_documentation_block:
```

seq_documentation_block:	854 bp	mRNA	ROD	13-APR-1999
LOCUS	AF097440	Mus musculus brain expressed x-linked protein 3 complete cds.		(13-APR-1999)
DEFINITION	AF097440	complete cds.		
ACCESSION	AF097440			
VERSION	AF097440.1	GI:4580593		
KEYWORDS		house mouse;		
SOURCE		MUS musculus		
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE		1. (bases 1 to 854)		
AUTHORS		Brown,A.L. and Kay,G.F.		
TITLE		Bexl, a gene with increased expression in parthenogenetic embryos, is a member of a novel gene family on the mouse X chromosome		
JOURNAL		Hum. Mol. Genet. 8 (4), 611-619 (1999)		
MEDLINE		99172070		
REMARK		Erratum: [[published erratum appears in Hum Mol Genet. 1999		

REFERENCE	May 78(5):943]]
AUTHORS	2 (bases 1 to 854)
TITLE	Brown, A.L. and Kay, G.F.
JOURNAL	Direct Submission
FEATURES	Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
SOURCE	Location/Qualifiers 1. 854

source

```

gene
1 tss_894 -type- pooled organism
CDS
17 gene_5246
17 gene_5246
/gene_5246
/gene_5246
/codon_start=1
/product=brain expressed x-
/protein_id=AA024431.1

```

```

/od_xref="GI:4580594"
/translation="MANVQENEMEQPLQNGGEDRPGGGEGHPAAANNNNHHNHN
HNHRRGQRIAPNFWAIPNQMNIDGLGGDDMMFMEEMREIRRLRELQLRNC
LRLMGELSNHHDDHDFECLMP"
237 a 212 c 228 g 177 t
BASE COUNT
ORIGIN

```

```

alignment_scores:
  Quality: 688.00      Length: 124
  Ratio: 5.548         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 98.387

alignment_block:
  US-09-327-750D-12 x AF097440

  Align seg 1/1 to: AF097440 from: 1 to: 854

1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
|||||
172 ATGGCCAATGTCCACCAGGAAAAACGAAGAGATGGAGCAGCCCTGCAGAA 221
|||||
17 nGlyGluGluAspArgProValGlyGlyGluGlyHisGlnProAlaG 34
|||||
222 TGGACAGGAAGACCGCCCTGTGGGAGGAGGTGAGGGCCACGAGCTGCTG 271
|||||
34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
|||||
272 CAACCAACAACAACAACAACCAACCAACCATACCAACAACCAACCAAGA 321
|||||
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||
322 GGCAGGCTCGCCGACCTTGCCCTTAACCTCCGATGGGCCATTCACCAAC 371
|||||
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
|||||
372 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATATGGAAATGTTC 421
|||||
84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
422 TGGAGAGAGATGAGAGAGATCCGGAAGAACGTTAGGGAGCTACAGCTGAG 471
|||||
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
|||||
472 AATGTGCTACGCATCCTTATGGGGGAGCTGCTAACCAACCAAGATCACCA 521
|||||
117 sAspGluPheCysLeuMetPro 124
|||||
522 TGATGAATTCTGCCTTATGCCT 543

```

seq_name: qb_ro:AF187065

seq documentation block:

seq_documentation_block:					
LOCUS	AF187065	519 bp	mRNA	ROD	11-JUN-2000
DEFINITION	Rattus norvegicus p75NTR-associated cell death executor (Nade) mRNA, complete cds.				

ACCESSION AF187065
 mRNA, complete cds.

ACCESSION	AF187063
VERSION	AF187065.1
	GI:8452895

KEYWORDS
REF ID: A66311 GI: 8432833

NETWORKS
SOURCE
Norway rat.

SOURCE	ORGANISM
NORWAY	<i>Rattus norvegicus</i>

ORGANISM
Rattus norvegicus
Eukaryota: Metazoa: Ch

Eukaryota; Metazoa; Chordata
Mammalia: Eutheria: Rodentia
Carnivora: Canidae: Felinae: Feline

Mammalia; Euthera; Rattus

REFERENCE 1 (bases 1 to 519)
Rattus.

REFERENCE
I (bases 1 to 519)
Mukai T Machine T
AUTHORS

AUTHORS
Mukai, J., Hachiya, T.,
Suzutani, D., Hasegawa, M.

Suvanto, P., Hanaoka, T.

TITLE
NADE, a p75NTR-associated

signal transduction me

p75NTR

J. Biol. Chem. 275 (23)

MEDLINE 20298829

REFERENCE 2 (bases 1 to 519)

AUTHORS Mukai, J., Hachiya, T.,

Hanaoka, T., Li, Y., Iri

TITLE
Direct Submission

JOURNAL
FILE
DATE
PRICE
SUBMISSION
Submitted (17-SEP-1999)

BOOKMAN
SUBMITTED (17 SEP-1993)
University, 630 West 1

FEATURES

FEATURES	LOCATION/QUAL
source	1. .519

SOURCE
T. 319

US-09-327-750d-12 x BG088461

Align seg 1/1 to: BG088461 from: 1 to: 599

```

1 MetAlaAsnValHisGlnGluMetGluGlnProLeuGlnAs 17
175 ATGGCCCAATGTCCACGAGAAACGAAGAGATGGAGCCCTGCAGAA 224
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlyHisGlnProAlaG 34
225 TGGCAGGAAGAACCCCTGTGGGAGGAGTGGAGGCCACCAAGCTGCTG 274
34 lyaAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
275 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 324
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrrPalaIleProAsnAr 67
325 GGCCAGGCTCGCCGACTTGCCTTAACCTCCGATGGGCCCATTCACCAACAG 374
67 gGlnMetAsnAspGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 84
375 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATGGAATGTTCA 424
84 etGluGluMetArgGluLeuArgLysLeuArgGluGluGluGluGluGluGlu 100
425 TGGAGGAGATGAGAGAGATCGGAGAAAGCTTACGGAGCTACAGCTGAGA 474
101 AsnCysLeuArgGluLeuMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 117
475 AATTGCTACGCATCCTTATGGGGAGCTGTCTAACCAACCAACGATCACA 524
117 sAspGluPheCysLeuMetPro 124
525 TGATGAATTCGCTTATGCGCT 546

```

seq_name: gb_est1:AA272375

seq_documentation_block:

LOCUS AA272375 612 bp mRNA EST 26-MAR-1997
DEFINITION vb69707.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
clone IMAGE:761628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AA272375
VERSION AA272375.1 GI:1910706
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 612)
Marta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:462548

Seq primer: -28ml3 rev2 Et from Amersham
High quality sequence stop: 507.
Location/Qualifiers
1. 612
/organism="Mus musculus"

/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:761628"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCAGCAATCTGAAGTGGAGCGCCCTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[GTGTGATTCGGTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead."

BASE COUNT 168 a 167 c 175 g 102 t
ORIGIN

alignment_scores:

Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750D-12 x AA272375

Align seg 1/1 to: AA272375 from: 1 to: 612

```

1 MetAlaAsnValHisGlnGluGluGluGluGluGluGluGluGluGlnAs 17
175 ATGGCCCAATGTCCACGAGAAACGAAGAGATGGAGCCCTGCAGAA 225
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlyHisGlnProAlaG 34
226 TGGACAGGAAGACCGCCTGTGGGAGGAGGTGAGGGCCACGAGCTGCTG 275
34 lyaAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
276 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 325
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrrPalaIleProAsnAr 67
326 GGCCAGGCTCGCCGACTTGCCTTAACCTCCGATGGGCCCATTCACCAACAG 375
67 gGlnMetAsnAspGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 84
376 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATGGAATGTTCA 425
84 etGluGluMetArgGluLeuArgLysLeuArgGluGluGluGluGluGluGlu 100
426 TGGAGGAGATGAGAGAGATCGGAGAAAGCTTACGGAGCTACAGCTGAGA 475
101 AsnCysLeuArgGluLeuMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 117
476 AATTGCTACGCATCCTTATGGGGAGCTGTCTAACCAACCAACGATCACA 525
117 sAspGluPheCysLeuMetPro 124
526 TGATGAATTCGCTTATGCGCT 547

```

seq_name: gb_est2:W81757

seq_documentation_block:

LOCUS W81757 616 bp mRNA EST 12-SEP-1996
DEFINITION me95d06.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
clone IMAGE:403307 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION W81757
VERSION W81757.1 GI:1392776
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 616)
Marra M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:247075

REFERENCE
AUTHORSTITLE
JOURNAL
COMMENT

Putative full length read
vector; to vector length is 617
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 492.
Location/Qualifiers
1..616
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:403307"
/clone_lib="Soares mouse embryo NDBE13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGGAGCGCGCGAAATTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru KO, Wayne
State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

FEATURES
source

seq_name: gb_est1:AM476468

seq_documentation_block:

LOCUS AM476468 628 bp mRNA EST 24-FEB-2000
DEFINITION uc76d02.y1 NCI-CGAP_Lu33 Mus musculus cDNA clone IMAGE:2937027 5'
similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
(HUMAN);, mRNA sequence.

ACCESSION AM476468
VERSION AM476468.1 GI:7046574

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 628)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: coapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

pva sequencing by: Washington University Genome Sequencing Center

pva distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium, .at:

www.ncbi.nlm.nih.gov/bbrp/image/image.htm

Seq primer: -47

Seq quality sequence stop: 456.

Location/Qualifiers

1..628

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:2937027"

/tissue_type="pooled lung tumors"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st

strand cDNA was prepared from mRNA obtained from pooled

lung tumors with a Not I - oligo(dT) primer [5'

TGTACCAATCTGAAGTGGGAGCGCGCGAAATTTTTTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES
source

BASE COUNT 159 a 159 c 159 g 139 t

alignment_scores:

Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750d-12 x W81757

Align seg 1/1 to: W81757 from: 1 to: 616

1 MetaAAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
|||||
7 ATGGCCAATGTCACCAAGAAACGAGATGGAGCGCCCTGCAGAA 56
|||||

17 nGlyGluGluAspArgProValGlyGlyGlyGluHisGlnProAlaG 34
|||||
57 TGACAGGAGGAGACCGCCCTGTGGGAGGAGGTGAGGCCACCGCTGCTG 106
|||||

34 LyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisHisArgArg 50
|||||
107 CAAAC 156
|||||

141 ATGGCCCAATGTCCACGAGAAACGAGAGATGGAGCCCTCGCAGAA 190
17 nGlyGluGluAspArgProValGlyGlyGlyGluGlyHisGlnProAlaG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
191 TGGACAGGAAGACCGCCCTGTGGGAGAGGTGAGGCCACACCGCTGCTG 240
34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 CAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 290
51 GlyGlnAlaArgGluLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
291 GGCAGGCTCGCGACTTCCCTTAACCTCCGATGGCCCATTCACCAACAG 340
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
341 GCAGATGAATGATGCGGTGGGAGATGGAGATGATATGGAAATGTTCA 390
84 eGluGluMetArgGluLeuLeuArgLysLeuArgGluLeuGlnLeuArg 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
391 TGGAGGAGATGAGAGATCCGAGAAAGCTTAGGAGCTACAGCTGAGA 440
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
441 AATTGCTACGCATCTTATGGGGAGCTGTCTAACCAACCAACCAACCA 490
117 sAspGluPheCysLeuMetPro 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
491 TGATGAATTCGCCCTATGCTT 512

seq_name: gb_est1:BE334877

seq_documentation_block:
LOCUS BE334877 585 bp mRNA EST 14-JUL-2000
DEFINITION US90b10.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:3325627.5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION BE334877
VERSION BE334877.1 GI:9208653
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 585)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1069791
Seq primer: -4ORP from Gibco
High quality sequence stop: 465.

FEATURES
SOURCE

1..585
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3325627"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 162 a 157 c 172 g 93 t 1 others
ORIGIN
alignment_scores:
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387
alignment_block:
US-09-327-750D-12 x BE334877
Align seg 1/1 to: BE334877 from: 1 to: 585

1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 ATGGCCAATGTCCACGAGAAACGAGAGATGGAGCCCTCGCAGAA 215
17 nGlyGluGluAspArgProValGlyGlyGlyGluGlyHisGlnProAlaG 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 TGGACAGGAAGACCGCCCTGTGGGAGAGGTGAGGCCACACCGCTGCTG 265
34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 CAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 315
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 GGCAGGCTCGCGACTTCCCTTAACCTCCGATGGGCCATTCACCAACAG 365
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
366 GCAGATGAATGATGCGGTGGGAGATGGAGATGATATGGAAATGTTCA 415
84 etGluGluMetArgGluLeuLeuArgArgLysLeuArgGluGluLeuArg 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 TGGAGGAGATGAGAGATCCGAGAAAGCTTAGGAGCTACAGCTGAGA 465
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 AATTGCTACGCATCTTATGGGGAGCTGTCTAACCAACCAACCAACCA 515
117 sAspGluPheCysLeuMetPro 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 TGATGAATTCGCCCTATGCTT 537

seq_name: gb_est2:W64711

seq_documentation_block:
LOCUS W64711 590 bp mRNA EST
DEFINITION md71g01.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
clone IMAGE:373872.5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION W64711
VERSION W64711.1 GI:1372353
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 590)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

10-JUN-1996

Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:235304
 Seq primer: ETPRIMER
 High quality sequence stop: 348.
 Location/Qualifiers
 1. 590
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_image="373872"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAATCTGAAGTGGAGCGCGCGGAATTTTATTTTATTTTATTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M. Fatima Bonaldo."
 BASE COUNT 165 a 155 c 166 g 104 t
 ORIGIN

alignment_scores:
 Quality: 688.00 Length: 124
 Ratio: 5.548 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.387
 alignment_block:
 US-09-327-750D-12 x W64711
 Align seg 1/1 to: W64711 from: 1 to: 590
 1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
 134 ATGCCCAATGTCACCAAGAAACGAAGAGATGGAGCCCTGCAGAA 183
 17 nGlyGluGluAspArgProValGlyGlyGlyHisGlnProAlaG 34
 184 TGGACAGGAAGACCGCCCTGTGGAGAGGTGTGGGCCACCGCTGCTG 233
 34 LyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArg 50
 234 CAACAAACAAACAAACAAACCAACCAACCAACCAACCAACCAAG 283
 51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaIleProAsn 67
 284 GCCCAGGCTGCCGACTTCCCTTCCGATGGGCCATTCCTCCCAACAG 333
 67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
 334 GCAGATGAATGACGGTGTGGTGGAGATGAGATGATATGGAATGTTCA 383
 84 eGluGluMetArgGluIleArgArgLysLeuArgGluLeuArg 100
 384 TGGAGAGATGAGAGATGCCGAGAAGCTTAGGAGCTACAGCTGAGA 433
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHis 117
 434 AATTGCTACGCATCTTATGGGAGCTGTCTAAACCAACCAACCAACCA 483
 117 sAspGluPheCysLeuMetPro 124

484 TGATGAATTCGCTTATGCT 505
 seq_name: gb_est2:BG088461

seq_documentation_block:
 LOCUS BG088461 599 bp mRNA EST 26-JAN-2001
 DEFINITION H3153D07-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3153D07 5', mRNA sequence.
 ACCESSION BG088461
 VERSION BG088461.1 GI:12571025
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 599)
 AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
 ,T.S., Carter,M.G. and Ko,M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)
 COMMENT Other_ESTs: H3153D07-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
 Plate: H3153 row: D column: 07
 Seq primer: -21M13 Reverse
 High quality sequence stop: 599
 POLYA-No.

FEATURES

Location/Qualifiers
 1. 599
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3153D07"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI. This
 clone is among a rearranged set of 15,247 clones from 11
 embryo cDNA libraries (including preimplantation stage
 embryos from unfertilized egg to blastocyst, embryonic
 part of E7.5 embryos, extraembryonic part of E7.5 embryos
 , and E12.5 female mesonephros/gonad) and one newborn
 ovary cDNA library. Average insert size 1.5 kb. All
 source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
 expression profiling of mid-gestation placenta and embryo
 using a 15,000 mouse developmental cDNA microarray, 2000,
 Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
 Large-scale cDNA analysis reveals phased gene expression
 patterns during preimplantation mouse development, 2000,
 Development, 127: 1737-1749; (3) Genome-wide mapping of
 unselected transcripts from extraembryonic tissue of
 7.5-day mouse embryos reveals enrichment in the t-complex
 and under-representation on the X chromosome, 1998, Hum
 Mol Genet 7: 1967-1978."
 BASE COUNT 163 a 163 c 176 g 97 t
 ORIGIN

alignment_scores:
 Quality: 688.00 Length: 124
 Ratio: 5.548 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.387
 alignment_block: